

*Durham University*

# Recent Advances in QCD Event Generators

Peter Richardson

IPPP, Durham University

# Introduction

- Monte Carlo event generators are essential for experimental particle physics.
- They are used for:
  - Comparison of experimental results with theoretical predictions;
  - Studies for future experiments.
- Often these programs are ignored by theorists and treated as black boxes by experimentalists.
- It is important to understand the assumptions and approximations involved in these simulations.

# Introduction

- Experimental physicists need to be able to answer the following questions
  - Is the effect I'm seeing due to different models, or approximations, or is it a bug?
  - Am I measuring a fundamental quantity or merely a parameter of the simulation code?
- Theorists need to understand enough to be able ask
  - Have the experimentalists misused the Monte Carlo giving incorrect results?

# Introduction

- For both the Tevatron and LHC we are interested in final states with large numbers of jets and leptons. For example
  - Top production
  - SUSY
- The backgrounds to these processes generally come from multiple QCD radiation giving jets.
- These QCD processes are of course interesting in their own right.



# Introduction

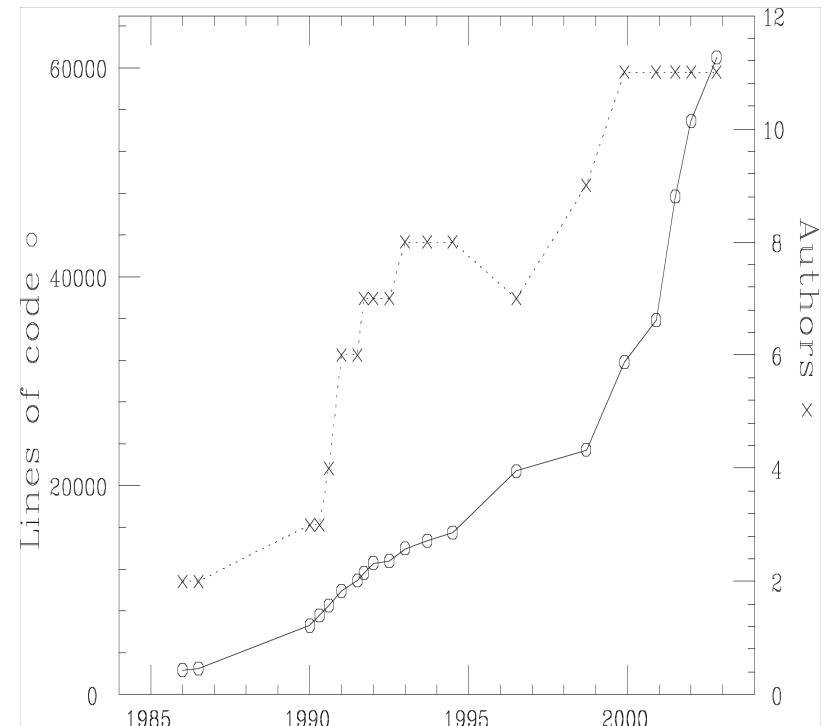
- In this talk I will start by describing the ideas behind Monte Carlo simulations.
- Recently there has been a lot of progress in two related areas:
  - Next-to-leading order simulation;
  - Matching leading order matrix elements;which are aimed at improving the treatment of hard radiation.
- I will go on to discuss these and where they are of use.

# Monte Carlo Event Generators

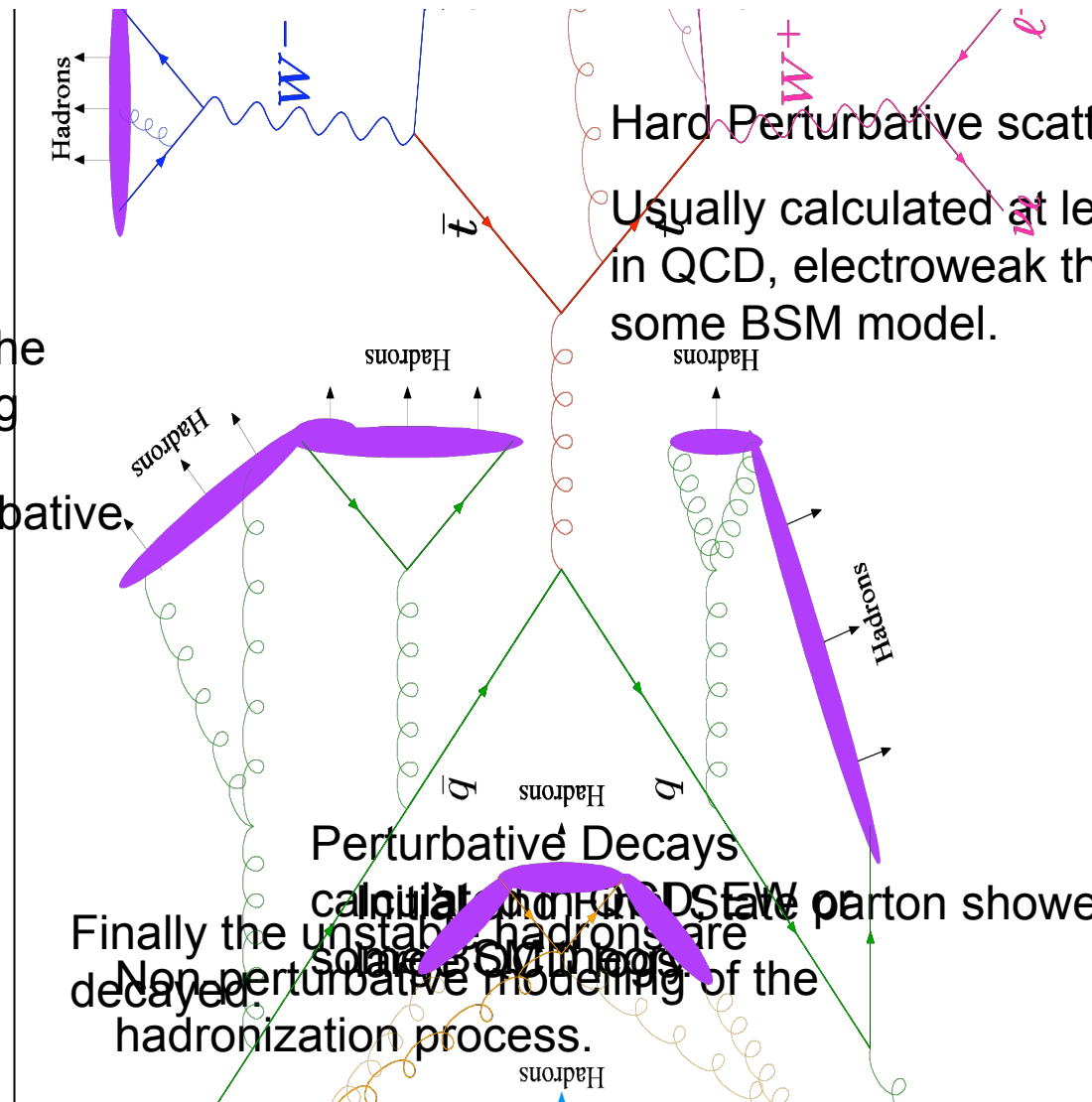
- There are a number of different Monte Carlo event generators in common use
  - ISAJET
  - PYTHIA
  - HERWIG
  - SHERPA
- They all split the event generation up into the same pieces.
- The models and approximations they use for the different pieces are of course different.

# C++ Generators

- Most of these programs are written in Fortran 77, (some are even older.)
- There are ongoing projects to rewrite HERWIG and PYTHIA in C++.
- Some of the newer projects, SHERPA, are also in C++.



# A Monte Carlo Event



Modelling of the soft underlying event  
Multiple perturbative scattering.

- Hard Perturbative scattering:
  - Usually calculated at leading order in QCD, electroweak theory or some BSM model.

Perturbative Decays  
 calculate the FSR state parton showers resum the  
 Finally the unstable hadrons are  
 Non-perturbative modeling of the  
 decayed.  
 hadronization process.

# Monte Carlo Event Generators

- All the event generators split the simulation up into the same phases:
  - Hard Process;
  - Parton Shower;
  - Secondary Decays;
  - Multiple Scattering/Soft Underlying Event;
  - Hadron Decays.
- I will briefly discuss the different models and approximations in the different programs.
- I will try and give a fair and objective comparison, but bear in mind that I'm one of the authors of **HERWIG**.

# QCD Radiation

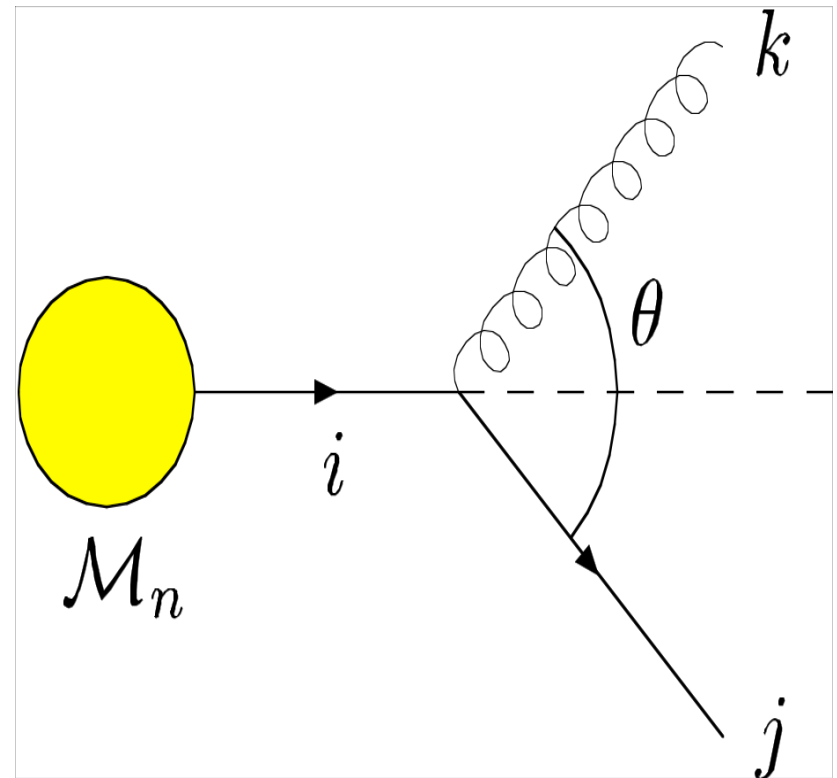
- It is impossible to calculate and **integrate** the matrix elements for large numbers of partons.
- Instead we treat the regions where the **emission** of QCD radiation is **enhanced**.
- This is **soft** and **collinear** radiation.
- The different generators differ in the sophistication of their simulation of this.

# Collinear Singularities

- In the **collinear** limit the cross section for a process **factorizes**

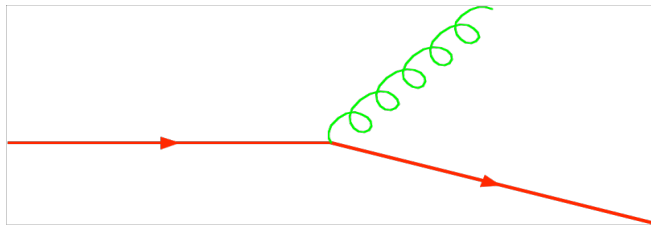
$$d\sigma_{n+1} = d\sigma_n \frac{d\theta^2}{\theta^2} dz \frac{\alpha_s}{2\pi} P_{ji}(z)$$

- $P_{ji}(z)$  is the DGLAP splitting function
- This expression is singular as  $\theta \rightarrow 0$ .
- What is a parton? (or what is the difference between a collinear pair and a parton)



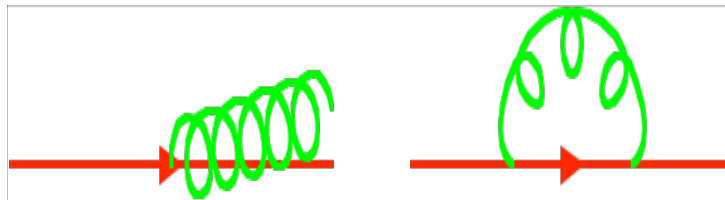
# Collinear Singularities

- Introduce a resolution criterion, e.g.  $k_T > Q_0$
- Combine the virtual corrections and unresolvable emission



Resolvable Emission

Finite



Unresolvable Emission

Finite

- Unitarity: Unresolved + Resolved = 1



# Monte Carlo Procedure

- Using this approach we can exponentiate the real emission piece.

$$\begin{aligned} \text{Unresolved} &= 1 - \text{Resolved} \\ &= 1 - \int_{q^2}^{Q^2} \frac{dk^2}{k^2} \int_{Q_0^2/q^2}^{1-Q_0^2/q^2} dz \frac{\alpha_s}{2\pi} P_{ji}(z) \\ &= \exp \left( - \int_{q^2}^{Q^2} \frac{dk^2}{k^2} \int_{Q_0^2/q^2}^{1-Q_0^2/q^2} dz \frac{\alpha_s}{2\pi} P_{ji}(z) \right) \end{aligned}$$

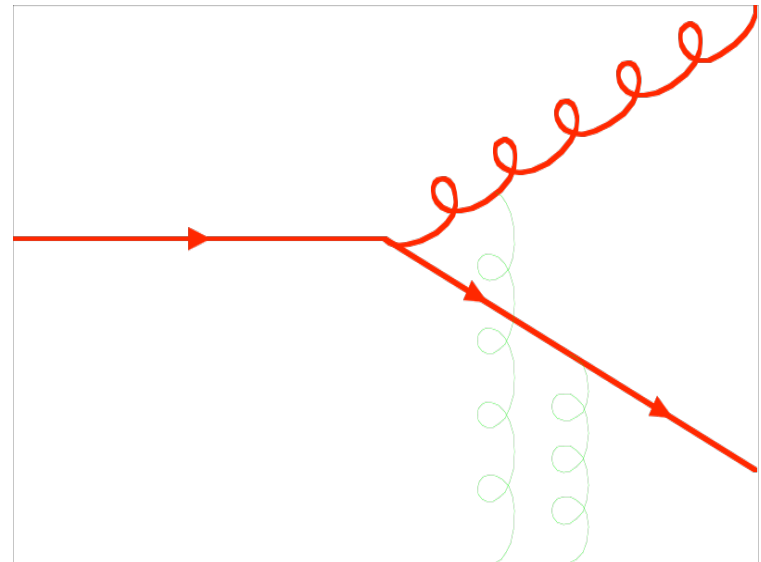
- This gives the **Sudakov form factor** which is the probability of evolving between two scales and emitting no resolvable radiation.
- More strictly it is the probability of evolving from a high scale to the cut-off with no resolvable emission.

# Monte Carlo Procedure

- The key difference between the different Monte Carlo simulations is in the choice of the evolution variable.
- Evolution Scale
  - Virtuality,  $q^2$
  - Transverse Momentum,  $k_T$ .
  - Angle,  $\theta$ .
  - ....
- Energy fraction,  $z$ 
  - Energy fraction
  - Light-cone momentum fraction
  - ....
- All are the same in the collinear limit.

# Soft Emission

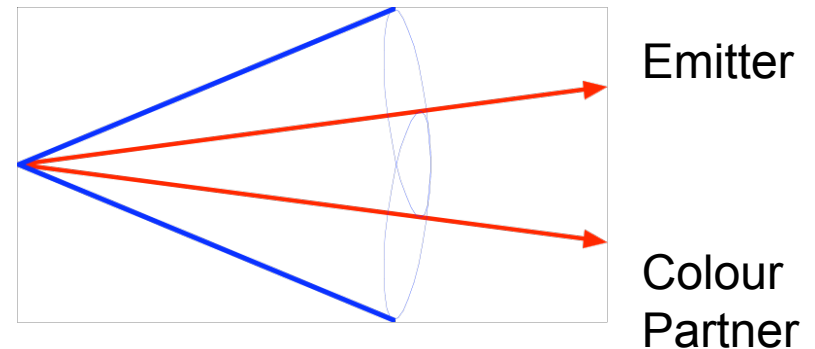
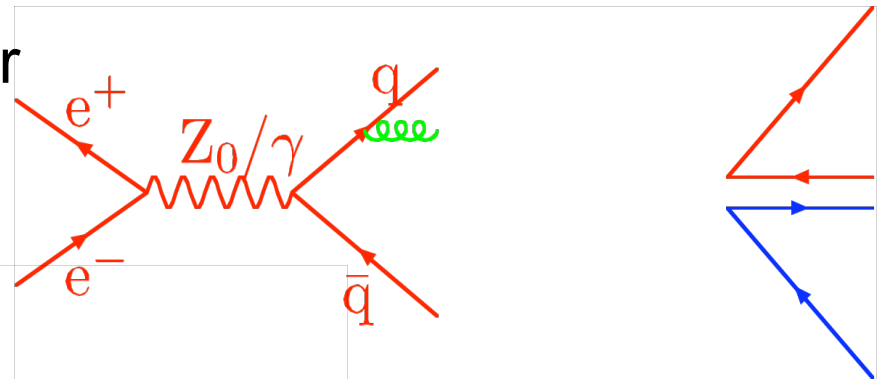
- However we have only considered collinear emission. What about soft emission?
- In the soft limit the matrix element factorizes but at the **amplitude** level.
- Soft gluons come from all over the event.
- There is quantum interference between them.
- Does this spoil the parton shower picture?



# Angular Ordering

Colour Flow

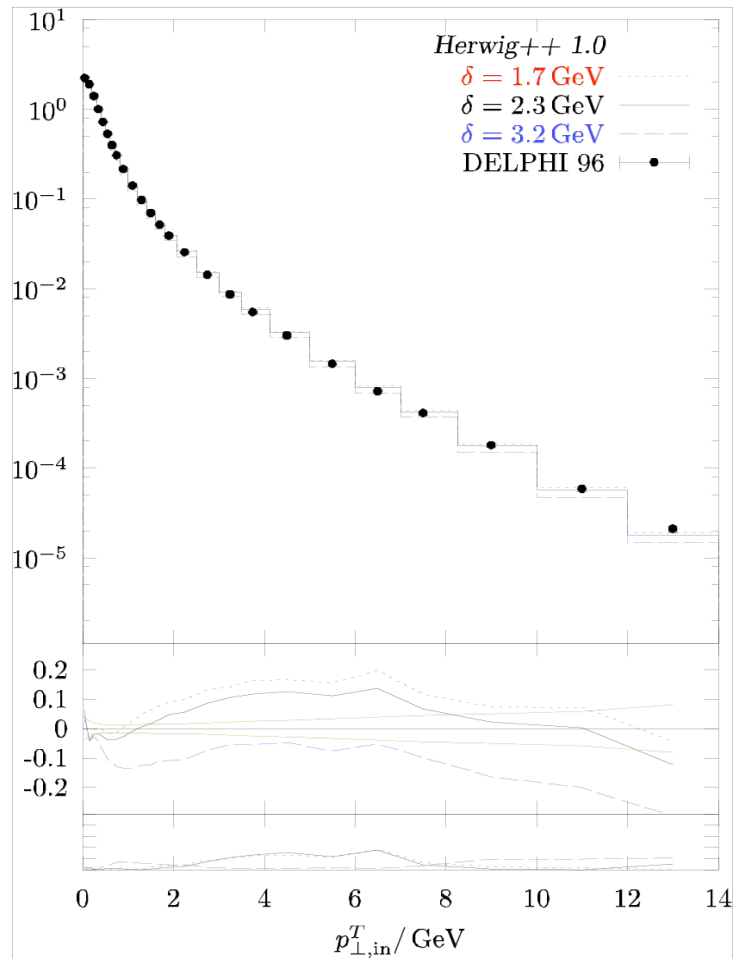
- There is a remarkable result that if we take the large number of colours limit much of the **interference** is **destructive**.
- In particular if we consider the colour flow in an event.
- QCD radiation only occurs in a cone up to the direction of the colour partner.
- The best choice of evolution variable is therefore an angular one.



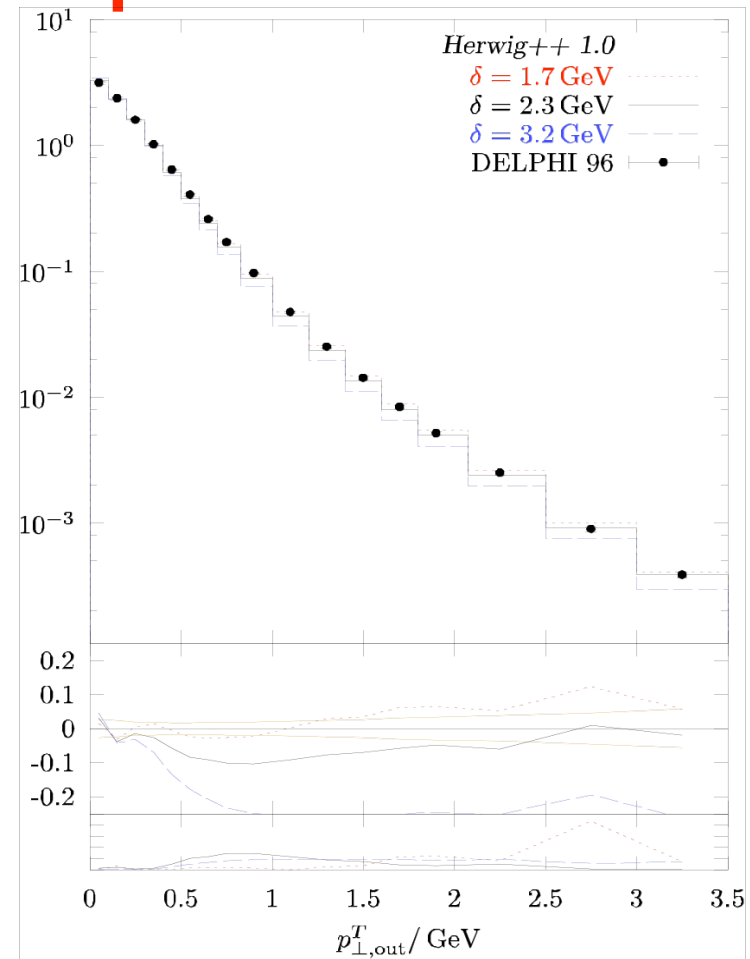
# Parton Shower

- **ISAJET** uses the original parton shower algorithm which only resums **collinear** logarithms.
- **HERWIG** uses the angular ordered parton shower algorithm which resums both **soft** and **collinear** singularities.
- **PYTHIA** uses the **collinear** algorithm with an angular **veto** to try to reproduce the effect of the angular ordered shower.
- **SHERPA** uses the **PYTHIA** algorithm.

# Event Shapes



Momentum transverse to the thrust axis in the event plane.



Momentum transverse to the thrust axis out of the event plane.

# Parton Shower

- The collinear algorithm implemented in **ISAJET** does not give good agreement with data.
- In general event generators which include angular ordering, colour coherence, give the best agreement with data.

# Dipole Showers

- The best agreement with the LEP data was obtained using **ARIADNE** which is based on the **dipole** approach.
- This is based on  $2 \rightarrow 3$  splittings rather than  $1 \rightarrow 2$  which makes it easier to conserve momentum.
- The soft and collinear are included in a consistent way.
- The initial state shower is more difficult in this approach though.

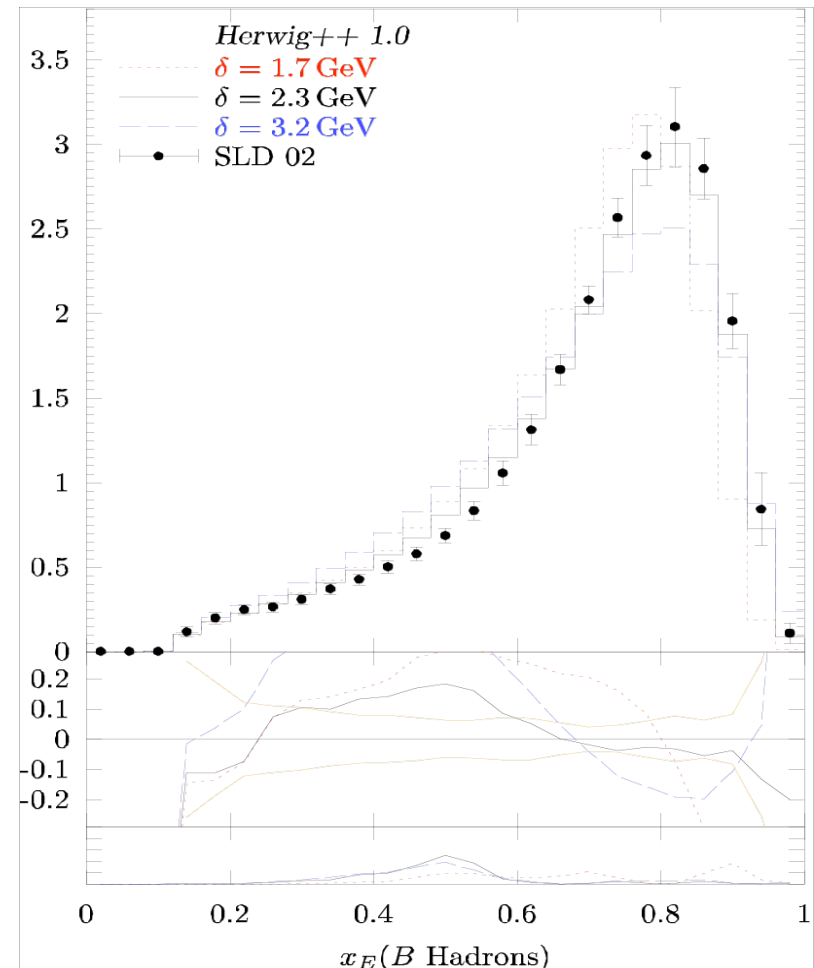


# Parton Showers

- Much of the recent work on parton showers has been on simulating hard radiation which I will talk about later.
- There are however some other **improvements**.
- The major new ideas are
  - An improved coherent parton shower using massive splitting functions.
  - A transverse momentum ordered shower.

# Herwig++ Shower

- Gieseke et. al.,  
JHEP 0402:005,2004  
JHEP 0312:045,2003.
- Gives an improved treatment of radiation from heavy particles, for example the b quark fragmentation function.
- This allows some radiation inside the 'dead-cone.'



# $P_T$ ordered shower

- T. Sjostrand hep-ph/0401061.
- Order the shower in transverse momentum rather than angle or virtuality.
- Still remains to shown that the coherence properties are correct.
- Can be used in new ideas in multiple scattering and the underlying event.
- T. Sjostrand, P.Z. Skands, hep-ph/0408302.

# Hadronization

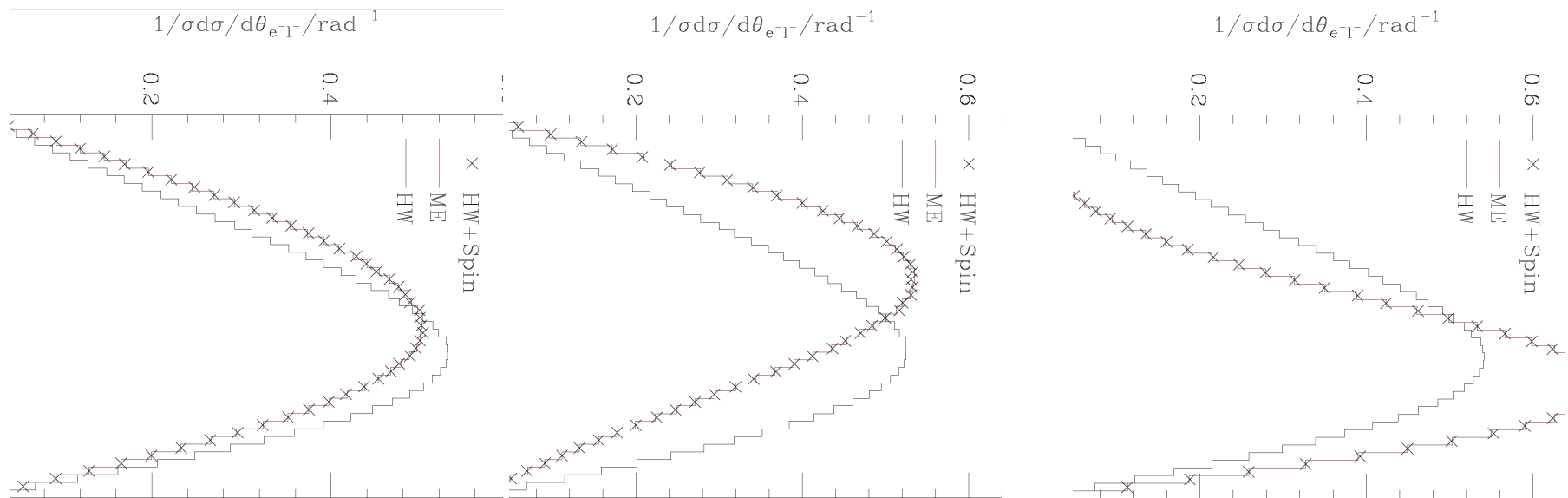
- As the hadronization is less important for what I will say later and there's been less progress I will only briefly mention the different models.
- **ISAJET** uses the original independent fragmentation model
- **PYTHIA** uses the Lund string model.
- **HERWIG** uses the cluster hadronization model.
- **ARIADNE** and **SHERPA** use the Lund model from PYTHIA.
- The independent fragmentation model cannot fit the LEP data.
- The cluster model gives good agreement with LEP data on event shapes but doesn't fit the identified particle spectrum as well.
- The Lund string model gives the best agreement with data.

# Signal Simulation

- In general we have become very **good** at simulating **signals**, be that top quark production, SUSY or other BSM physics.
- In many cases the simulations, particularly in HERWIG, the simulation is very **detailed** including **correlation** effects.
- This should be good enough for top and is certainly good enough for things that haven't been seen yet.

# Signal Simulation

Angle between the lepton in top decay and the beam for top pair production at a 500 GeV linear collider.



# Hard Jet Radiation

- I've tried to show you that the parton shower is designed to simulate **soft** and **collinear** radiation.
- While this is the bulk of the emission we are often interested in the radiation of a **hard** jet.
- This is **not** something the **parton shower** should be able to do, although it often does better than we expect.
- **If you are looking at hard radiation HERWIG and PYTHIA will often get it wrong.**

# Hard Jet Radiation

- Given this obvious failing of the approximations this is an obvious area to make improvements in the shower and has a long history.
- You will often here this called
  - Matrix Element matching.
  - Matrix Element corrections.
  - Merging matrix elements and parton shower
  - MC@NLO
- I will discuss all of these and where the different ideas are useful.

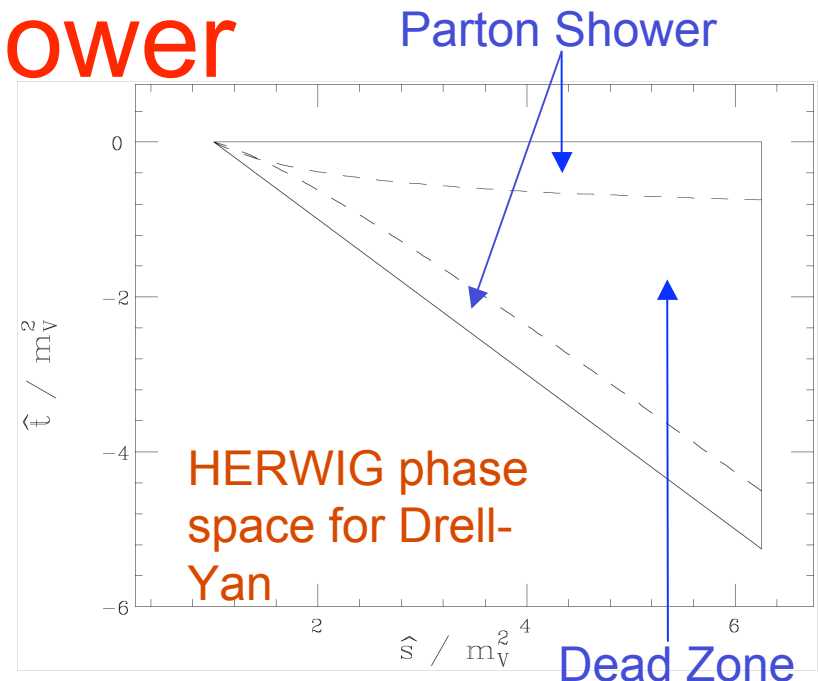


# Hard Jet Radiation: General Idea

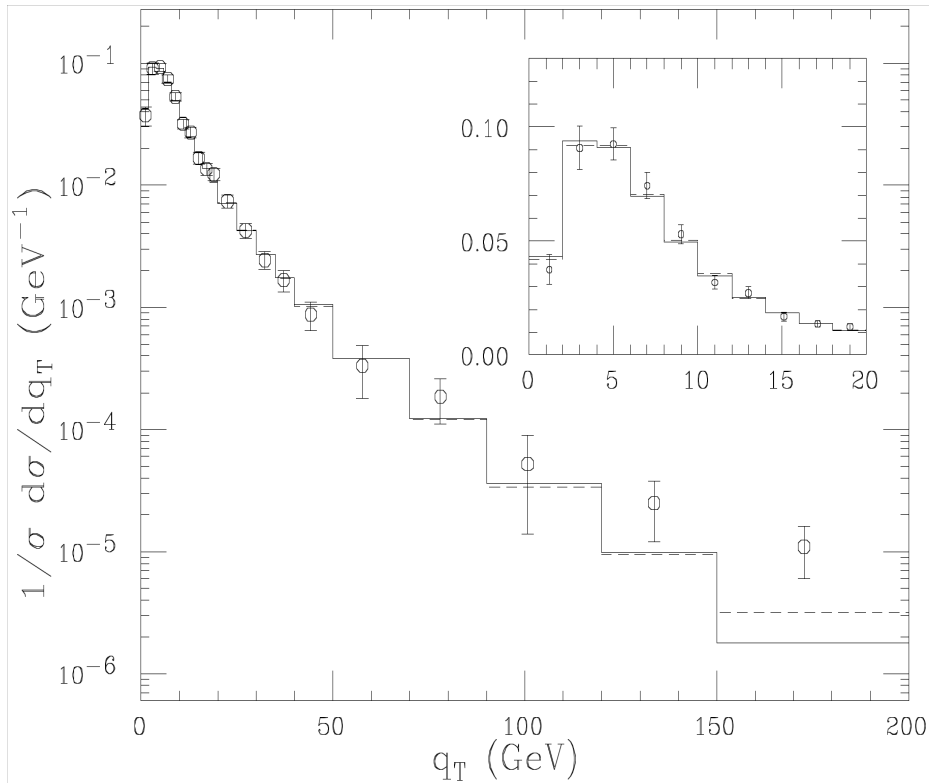
- **Parton Shower** (PS) simulations use the soft/collinear approximation:
  - Good for simulating the internal structure of a jet;
  - Can't produce high  $p_T$  jets.
- **Matrix Elements** (ME) compute the exact result at fixed order:
  - Good for simulating a few high  $p_T$  jets;
  - Can't give the structure of a jet.
- We want to use both in a **consistent** way, i.e.
  - ME gives hard emission
  - PS gives soft/collinear emission
  - Smooth matching between the two.
  - No double counting of radiation.

# Matching Matrix Elements and Parton Shower

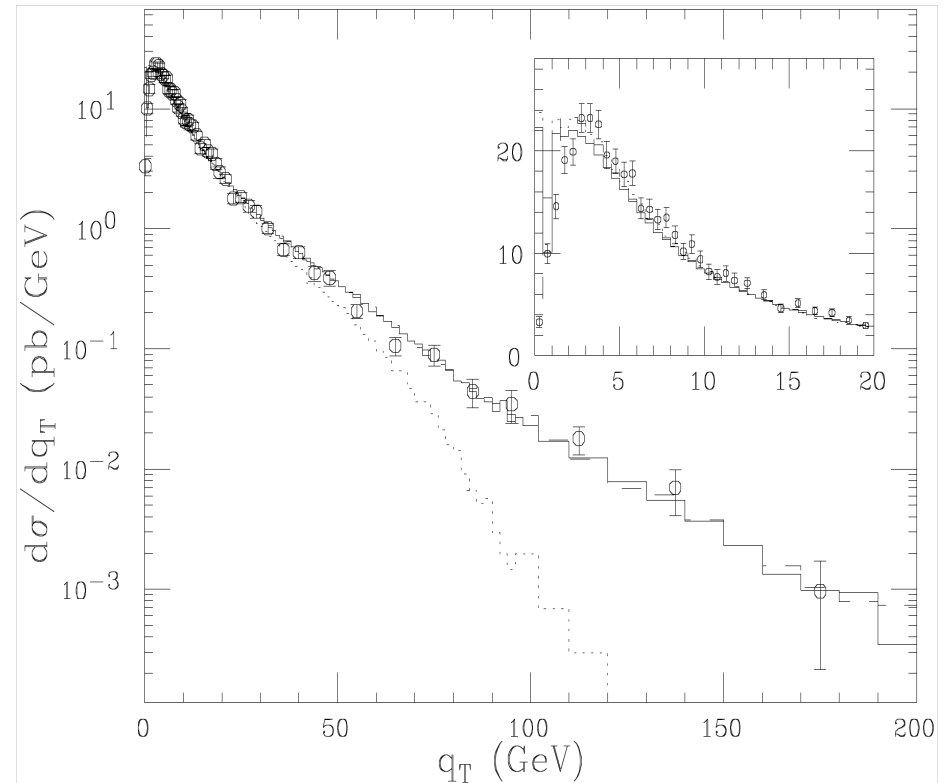
- The oldest approaches are usually called matching matrix elements and parton showers or the **matrix element correction**.
- Slightly different for **HERWIG** and **PYTHIA**.
- In **HERWIG**
  - Use the leading order matrix element to fill the dead zone.
  - Correct the parton shower to get the leading order matrix element in the already filled region.
- PYTHIA fills the full phase space so only the second step is needed.



# Matrix Element Corrections



W  $q_T$  distribution from D0



Z  $q_T$  distribution from CDF

G. Corcella and M. Seymour, Nucl.Phys.B565:227-244,2000.

# Matrix Element Corrections

- There was a lot of work for both HERWIG and PYTHIA and the corrections for
  - $e^+e^-$  to hadrons
  - DIS
  - Drell-Yan
  - Top Decay
  - Higgs Production
- There are problems with this
  - Only the hardest emission was correctly described
  - The leading order normalization was retained.

# Recent Progress

- In the last few years there has been a lot of work addressing both of these problems.
- Two types of approach have emerged

## 1) NLO Simulation

- NLO normalization of the cross section
- Gets the hardest emission correct

## 2) Multi-Jet Leading Order

- Still leading order.
- Gets many hard emission correct.

# NLO Simulation

- There has been a lot of work on NLO Monte Carlo simulations.
- However apart from some early work by Dobbs the only Frixione, Nason and Webber have produced code which can be used to generate results.
- I will therefore only talk about the work of Frixione, Nason and Webber.
- Most of this is taken from Bryan Webber's talk at the YETI meeting in Durham.

# MC@NLO

- S. Frixione and B.R. Webber JHEP 0206(2002) 029, hep-ph/0204244, hep-ph/0309186
- S. Frixione, P. Nason and B.R. Webber, JHEP 0308(2003) 007, hep-ph/0305252.
- <http://www.hep.phy.cam.ac.uk/theory/webber/MCatNLO/>

# MC@NLO

- MC@NLO was designed to have the following features.
  - The output is a set of fully exclusive events.
  - The total rate is accurate to NLO
  - NLO results for observables are recovered when expanded in  $\alpha_s$ .
  - Hard emissions are treated as in NLO calculations.
  - Soft/Collinear emission are treated as in the parton shower.
  - The matching between hard emission and the parton shower is smooth.
  - MC hadronization models are used.



# Toy Model

- I will start with Bryan Webber's **toy model** to explain MC@NLO to discuss the **key features** of NLO, MC and the matching.
- Consider a system which can radiate photons with energy  $x$  with

$$0 \leq x \leq x_s \leq 1$$

where  $x_s$  is the energy of the system before radiation.

- After radiation the energy of the system  $x'_s = x_s - x$
- Further radiation is possible but photons don't radiate.

# Toy Model

- Calculating an observable at NLO gives

$$\langle O \rangle = \lim_{\varepsilon \rightarrow 0} \int_0^1 dx x^{-2\varepsilon} O(x) \left[ \left( \frac{d\sigma}{dx} \right)_B + \left( \frac{d\sigma}{dx} \right)_V + \left( \frac{d\sigma}{dx} \right)_R \right]$$

where the Born, Virtual and Real contributions are

$$\left( \frac{d\sigma}{dx} \right)_B = B \delta(x) \quad \left( \frac{d\sigma}{dx} \right)_V = \alpha \left( \frac{B}{2\varepsilon} + V \right) \delta(x) \quad \left( \frac{d\sigma}{dx} \right)_R = \alpha \frac{R(x)}{x}$$

$\alpha$  is the coupling constant and  $\lim_{x \rightarrow 0} R(x) = B$

# Toy Model

- In a subtraction method the real contribution is written as

$$\langle O \rangle_R = \alpha B O(0) \int_0^1 dx \frac{1}{x^{1+2\varepsilon}} + \alpha \int_0^1 dx \frac{O(x)R(x) - BO(0)}{x^{1+2\varepsilon}}$$

- The second integral is finite so we can set  $\varepsilon = 0$

$$\langle O \rangle_R = -\alpha \frac{B}{2\varepsilon} O(0) + \alpha \int_0^1 dx \frac{O(x)R(x) - BO(0)}{x}$$

- The NLO prediction is therefore

$$\langle O \rangle_{sub} = \int_0^1 dx \left[ \alpha O(x) \frac{R(x)}{x} + O(0) \left( B + \alpha V - \alpha \frac{B}{x} \right) \right]$$

# Toy Monte Carlo

- In a MC treatment the system can emit many photons with the probability controlled by the Sudakov form factor, defined here as

$$\Delta(x_1, x_2) = \exp\left[-\alpha \int_{x_1}^{x_2} dx \frac{Q(x)}{x}\right]$$

where  $Q(x)$  is a monotonic function which has

$$0 \leq Q(x) \leq 1 \qquad \lim_{x \rightarrow 0} Q(x) = 1 \qquad \lim_{x \rightarrow 1} Q(x) = 0$$

- $\Delta(x_1, x_2)$  is the probability that no photon can be emitted with energy  $x$  such that  $x_1 \leq x \leq x_2$  .

# Toy MC@NLO

- We want to interface NLO to MC. Naïve first try
  - $O(0) \Rightarrow$  start MC with 0 real emissions:  $F_{MC}^0$
  - $O(x) \Rightarrow$  start MC with 1 real emission at  $x$ :  $F_{MC}^1(x)$
- So that the overall generating functional is

$$\int_0^1 dx \left[ F_{MC}^0 \left( B + \alpha V - \frac{\alpha B}{x} \right) + F_{MC}^1(x) \frac{\alpha R(x)}{x} \right]$$

- This is wrong because MC with no emissions will generate emission with NLO distribution

$$\left( \frac{d\sigma}{dx} \right)_{MC} = \alpha B \frac{Q(x)}{x}$$

# Toy MC@NLO

- We must subtract this from the second term

$$F_{MC@NLO} = \int_0^1 dx \left[ F_{MC}^0 \left( B + \alpha V + \frac{\alpha B(Q(x) - 1)}{x} \right) + F_{MC}^1(x) \frac{\alpha(R(x) - BQ(x))}{x} \right]$$

- This prescription has many good features:
  - The added and subtracted terms are equal to  $O(\alpha)$
  - The coefficients of  $F_{MC}^0$  and  $F_{MC}^1$  are separately finite.
  - The resummation of large logs is the same as for the Monte Carlo renormalized to the correct NLO cross section.

However some events may have negative weight.

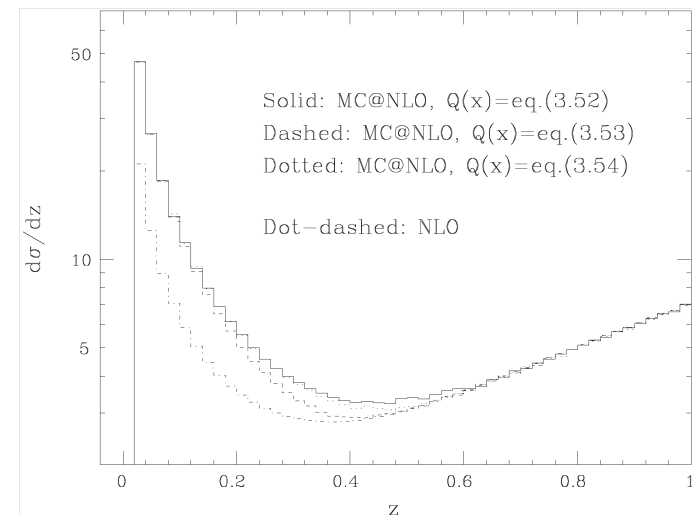
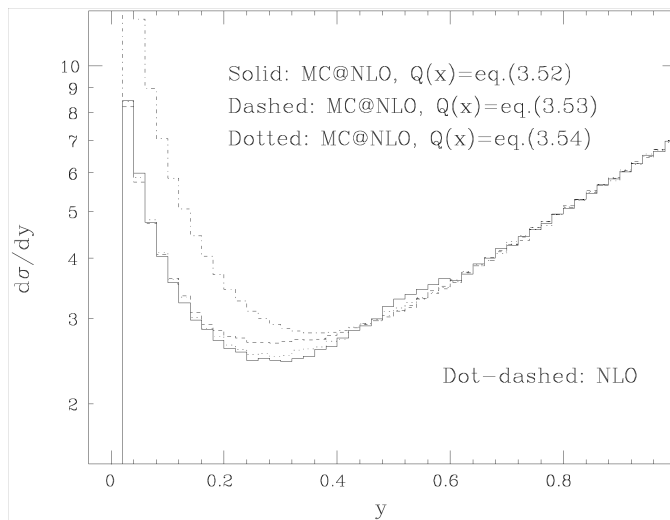
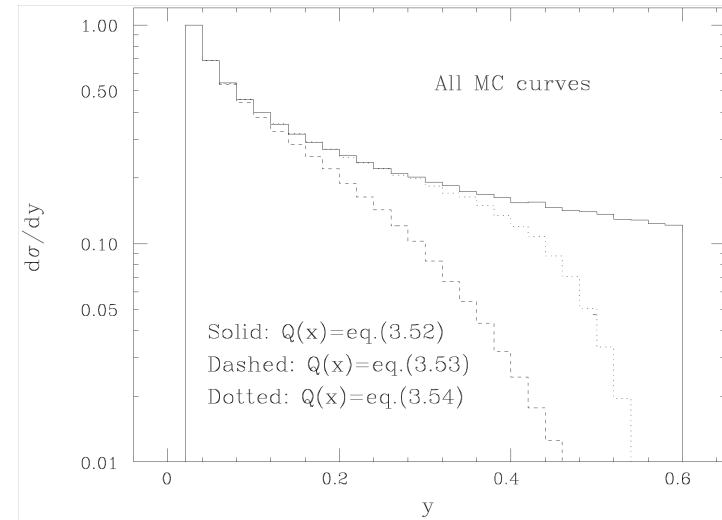
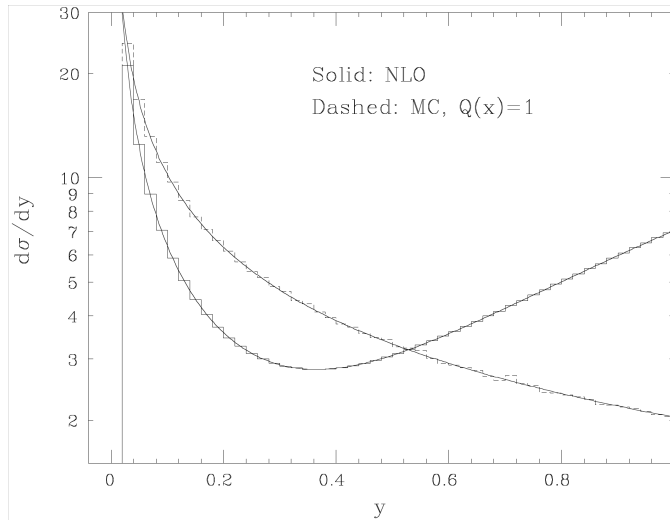
# Toy MC@NLO Observables

- As an example of an “exclusive” observable consider the energy  $y$  of the hardest photon in each event.
- As an “inclusive” observable consider the fully inclusive distributions of photon energies,  $z$
- Toy model results shown are for

$$\alpha = 0.3, \quad B = 2, \quad V = 1,$$

$$R(x) = B + x \left( 1 + \frac{x}{2} + 20x^2 \right)$$

# Toy MC@NLO Observables





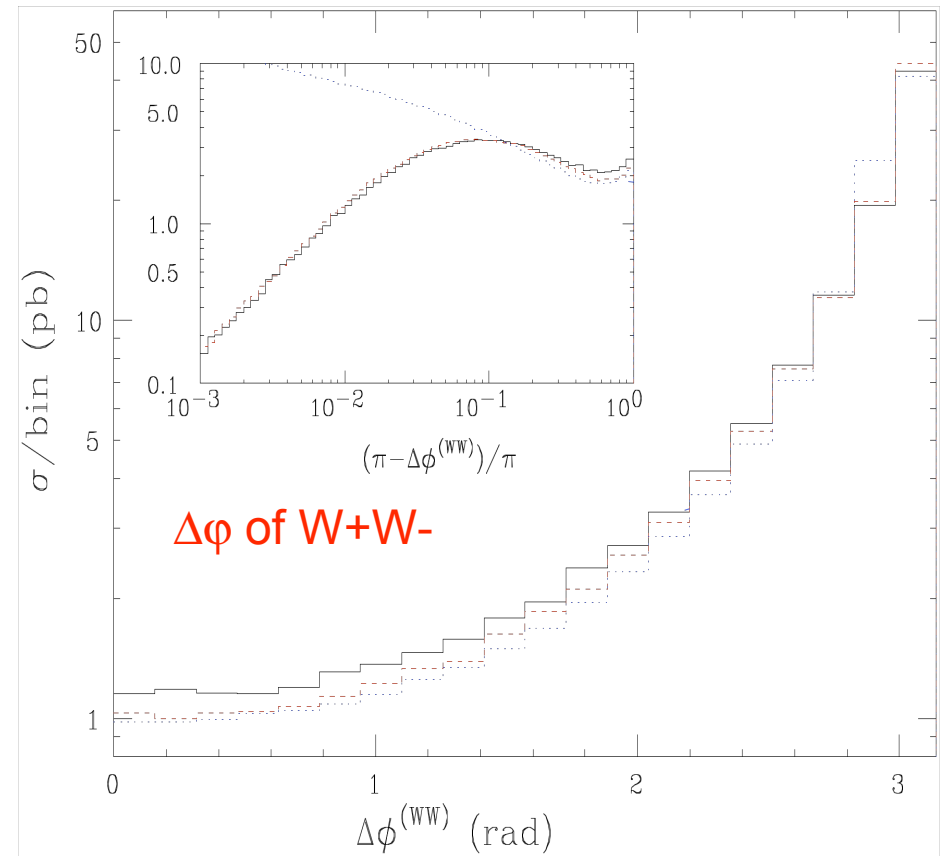
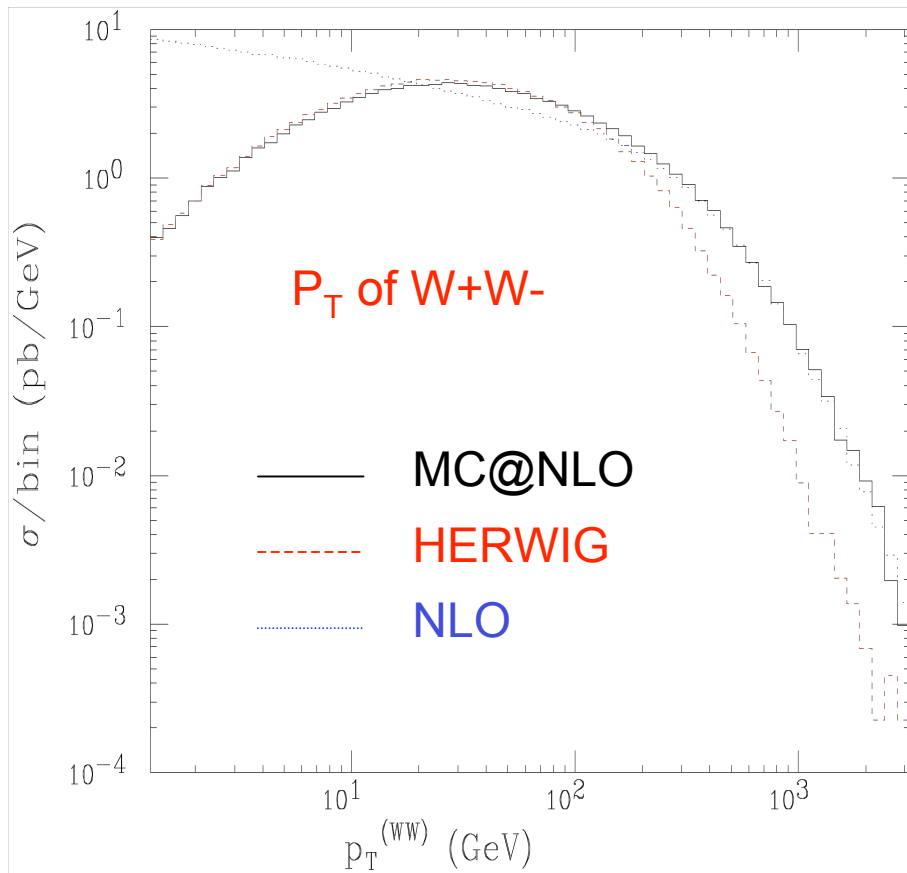
# Real QCD

- For **normal QCD** the **principle** is the **same** we subtract the shower approximation to the real emission and add it to the virtual piece.
- This cancels the singularities and avoids double counting.
- It's a lot more complicated.

# Real QCD

- For each new process the shower approximation must be worked out, which is often complicated.
- While the general approach works for any shower it has to be worked out for a specific case.
- So for MC@NLO only works with the HERWIG shower algorithm.
- It could be worked out for PYTHIA or Herwig++ but this remains to be done.

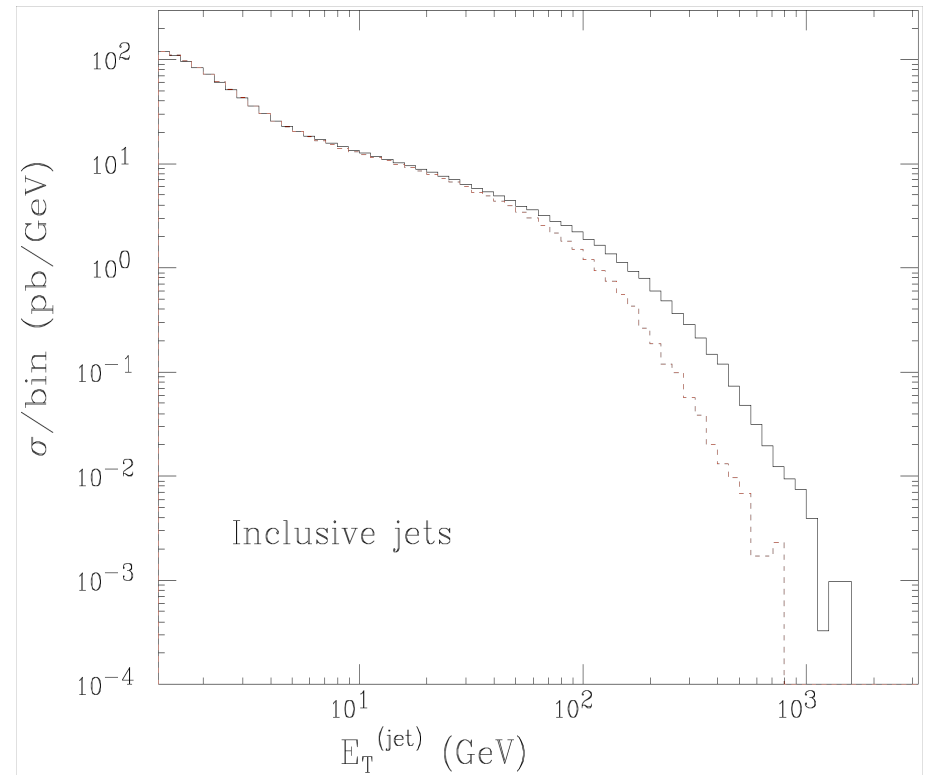
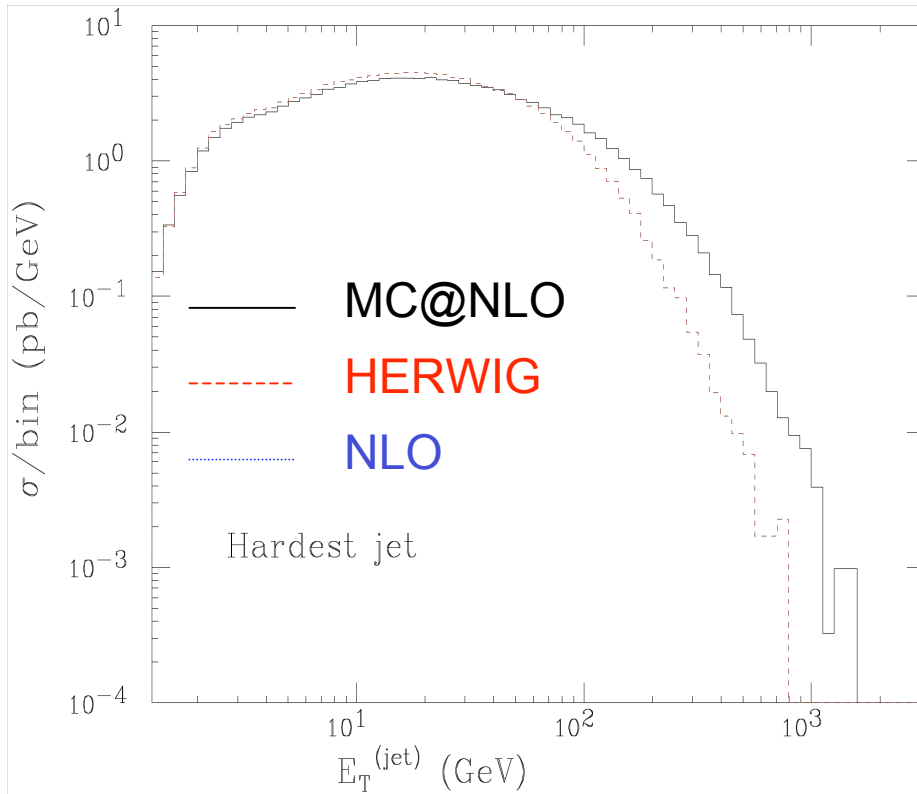
# $W^+W^-$ Observables



MC@NLO gives the correct high  $P_T$  result and soft resummation.

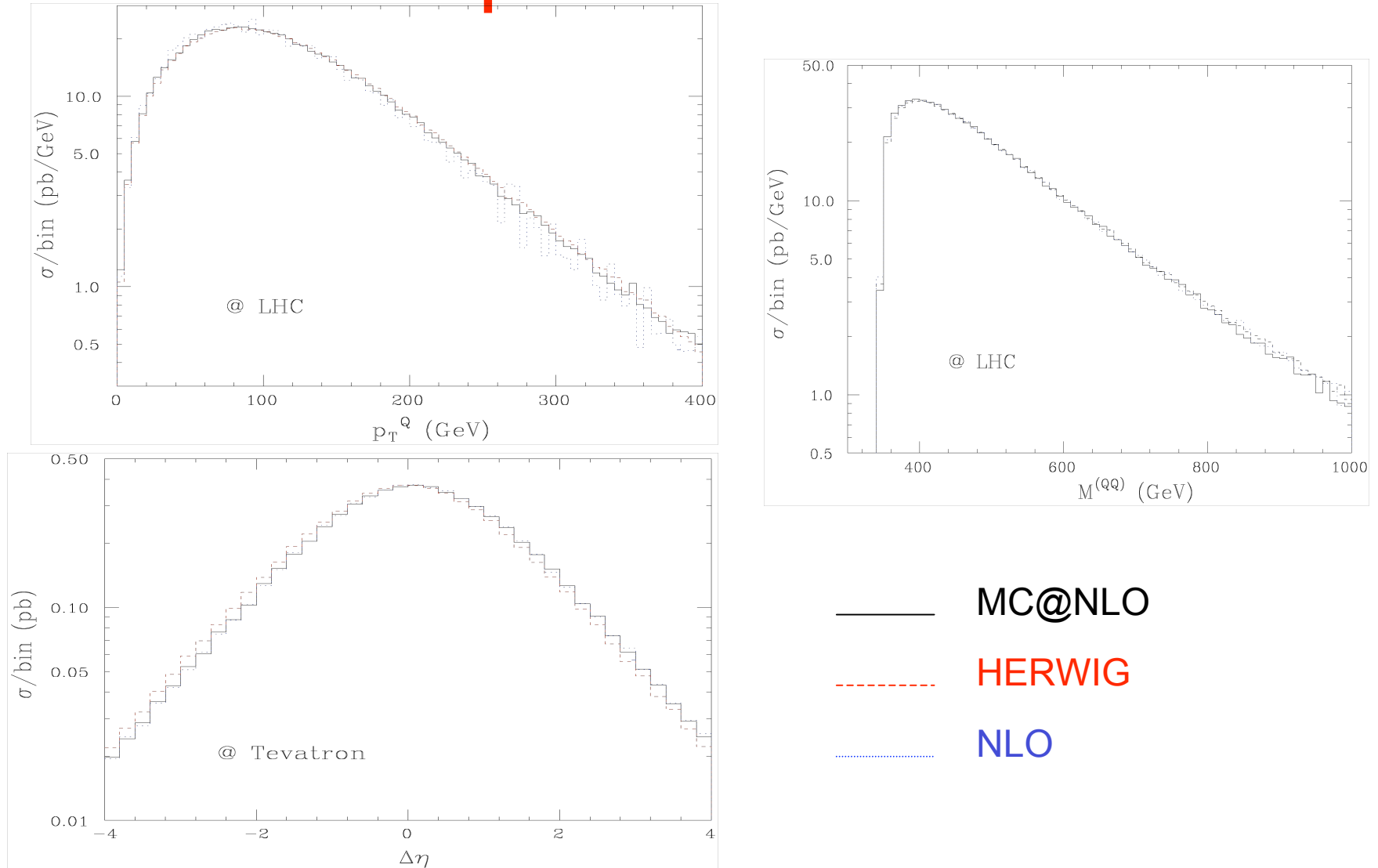
S. Frixione and B.R. Webber JHEP 0206(2002) 029, hep-ph/0204244, hep-ph/0309186

# $W^+W^-$ Jet Observables



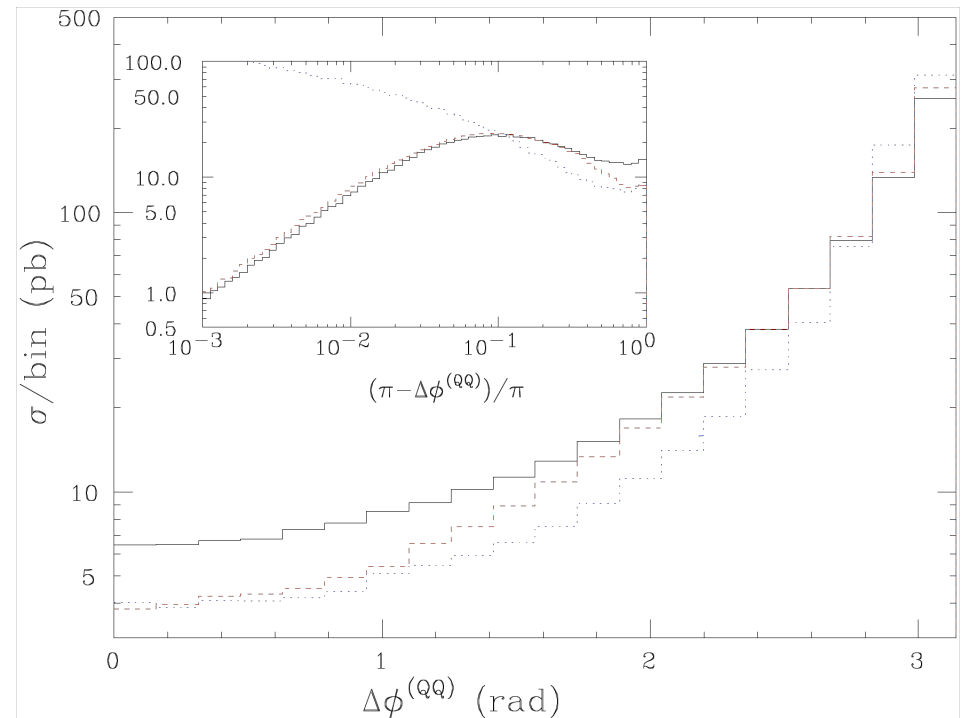
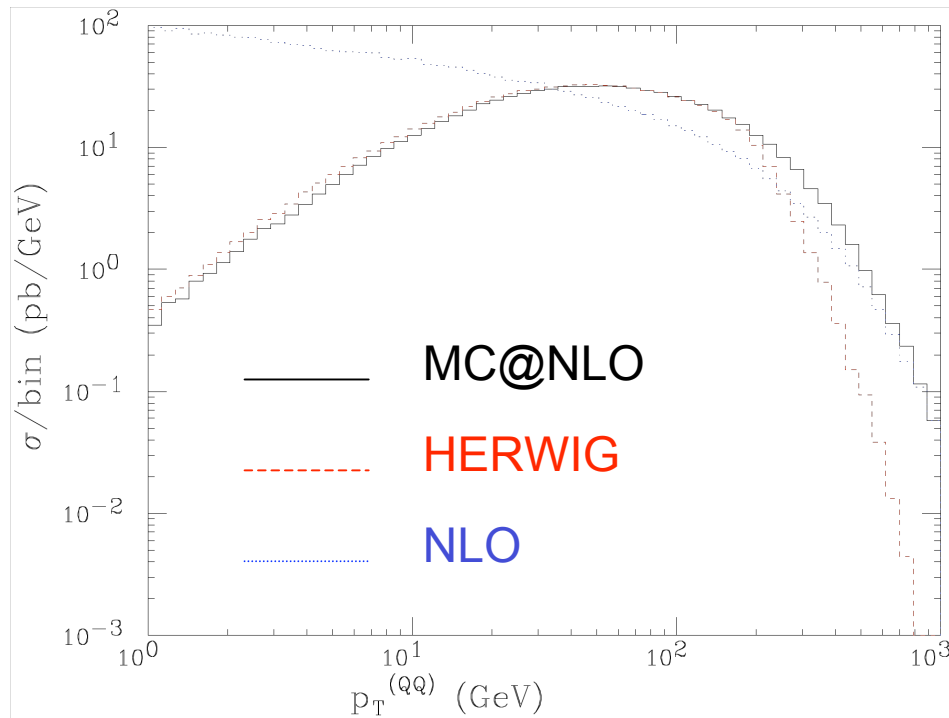
S. Frixione and B.R. Webber JHEP 0206(2002) 029, hep-ph/0204244, hep-ph/0309186

# Top Production



S. Frixione, P. Nason and B.R. Webber, JHEP 0308(2003) 007, hep-ph/0305252.

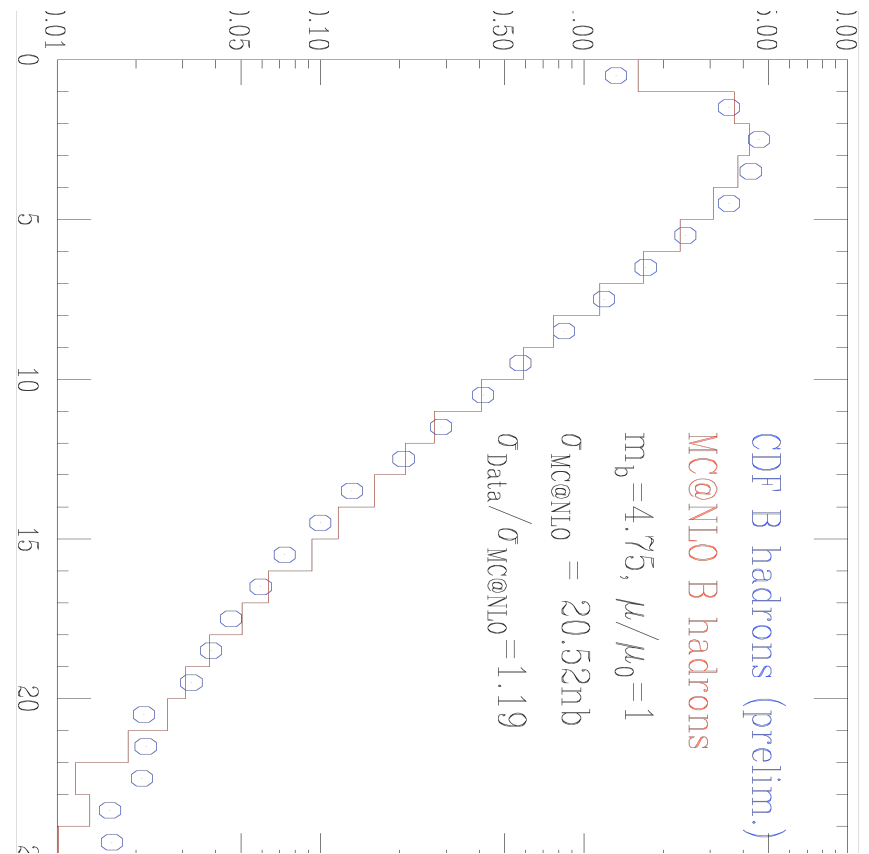
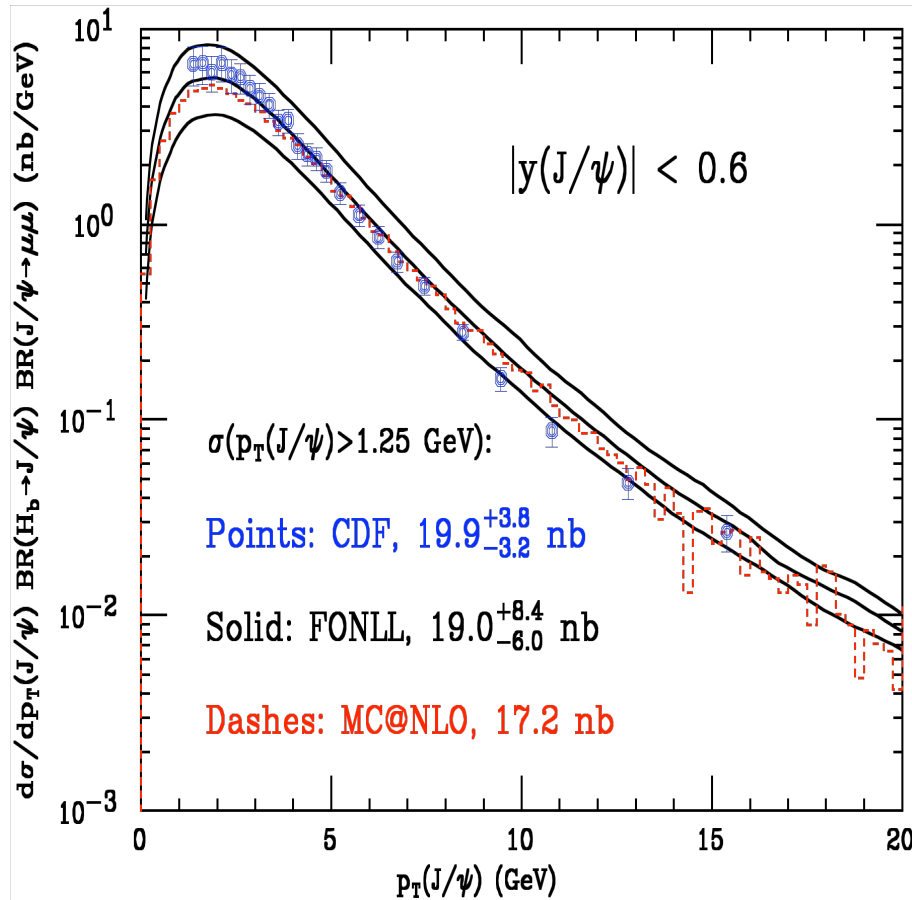
# Top Production at the LHC



S. Frixione, P. Nason and B.R. Webber, JHEP 0308(2003) 007, hep-ph/0305252.

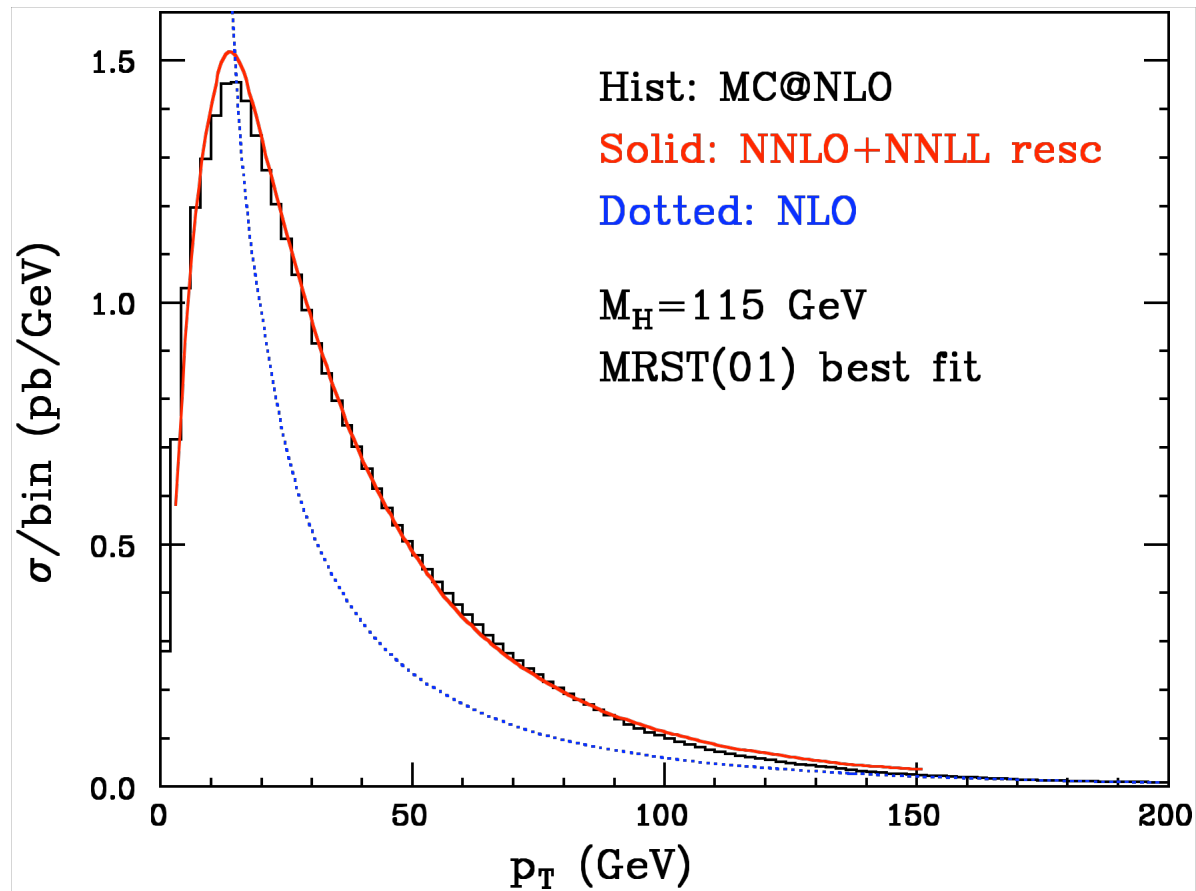
Bonn Seminar 27<sup>th</sup> January

# B Production at the Tevatron



S. Frixione, P. Nason and B.R. Webber, JHEP 0308(2003) 007, hep-ph/0305252.

# Higgs Production at LHC



S. Frixione and B.R. Webber JHEP 0206(2002) 029, hep-ph/0204244, hep-ph/0309186



# NLO Simulation

- So far MC@NLO is the only implementation of a NLO Monte Carlo simulation.
- Recently there have been some ideas by Paulo Nason JHEP 0411:040,2004.
- Here there would be no negative weights but more terms would be exponentiated beyond leading log.
- This could be an improvement but we will need to see physical results.

# Multi-Jet Leading Order

- While the **NLO** approach is good for **one hard** additional jet and the overall **normalization** it **cannot** be used to give **many jets**.
- Therefore to simulate these processes use matching at **leading order** to get many hard emissions correct.
- I will briefly review the general idea behind this approach and then show some results.

# CKKW Procedure

- Catani, Krauss, Kuhn and Webber JHEP 0111:063,2001.
- In order to match the ME and PS we need to separate the phase space:
- One region contains the soft/collinear region and is filled by the PS;
- The other is filled by the matrix element.
- In these approaches the phase space is separated using in  $k_T$ -type jet algorithm.

# Durham Jet Algorithm

- For all final-state particles compute the resolution variables

$$d_{kB} \approx E_k^2 \theta_{kB}^2 \approx k_{\perp kB}^2 \quad \theta_{kB}^2 \rightarrow 0$$

$$d_{kl} \approx \min(E_k^2, E_l^2) \theta_{kl}^2 \approx k_{\perp kl}^2 \quad \theta_{kl}^2 \rightarrow 0$$

- The smallest of these is selected. If  $d_{kl}$  is the smallest the two particles are merged. If  $d_{kB}$  is the smallest the particle is merged with the beam.
- This procedure is repeated until the minimum value is above some stopping parameter  $d_{\text{cut}}$ .
- The remaining particles and pseudo-particles are then the hard jets.

# CKKW Procedure

- Radiation above a cut-off value of the jet measure is simulated by the matrix element and radiation below the cut-off by the parton shower.

- 1) Select the jet multiplicity with probability

$$P_n = \frac{\sigma_n}{\sum_{k=0}^N \sigma_k}$$

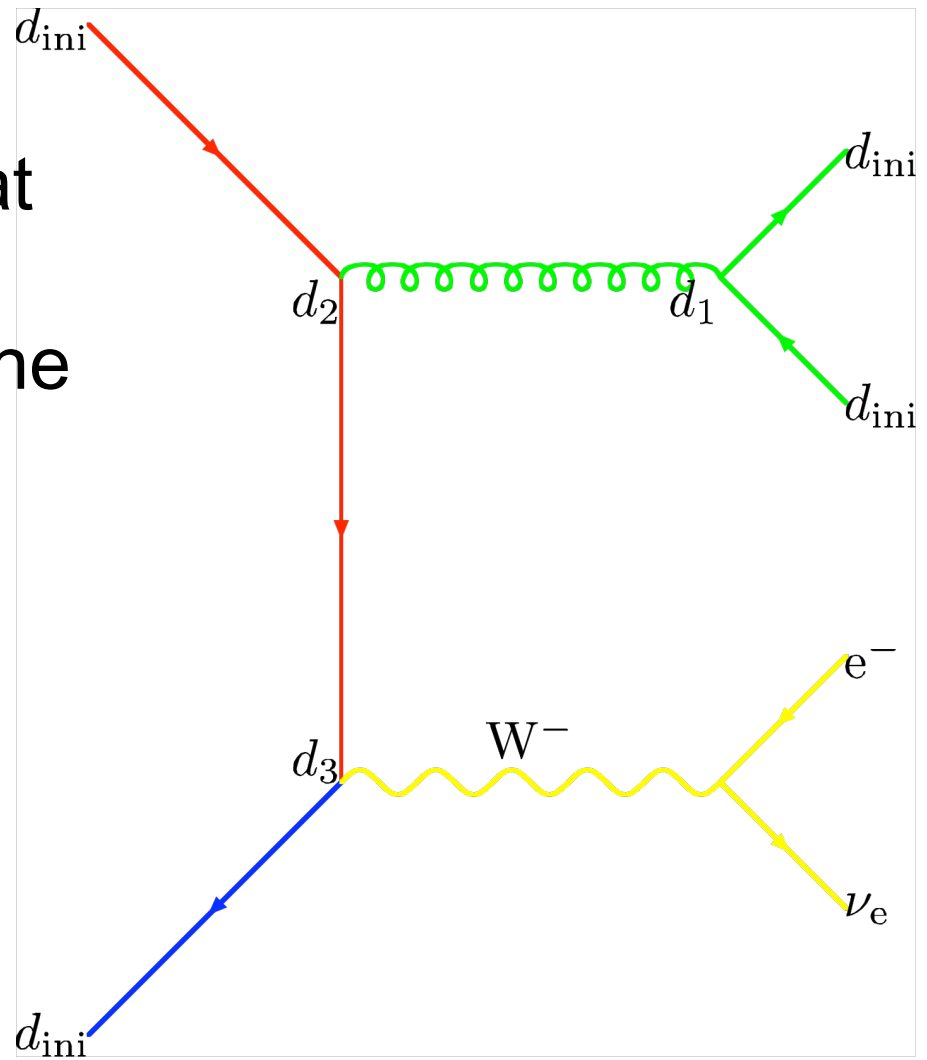
where  $\sigma_n$  is the  $n$ -jet matrix element evaluated at resolution  $d_{\text{ini}}$  using  $d_{\text{ini}}$  as the scale for the PDFs and  $\alpha_S$ ,  $n$  is the jet of jets

- 2) Distribute the jet momenta according the ME.

# CKKW Procedure

- 3) Cluster the partons to determine the values at which 1,2,.. $n$ -jets are resolved. These give the nodal scales for a tree diagram.
- 4) Apply a coupling constant reweighting.

$$\frac{\alpha_S(d_1)\alpha_S(d_2)...\alpha_S(d_3)}{\alpha_S(d_{\text{ini}})^n} \leq 1$$

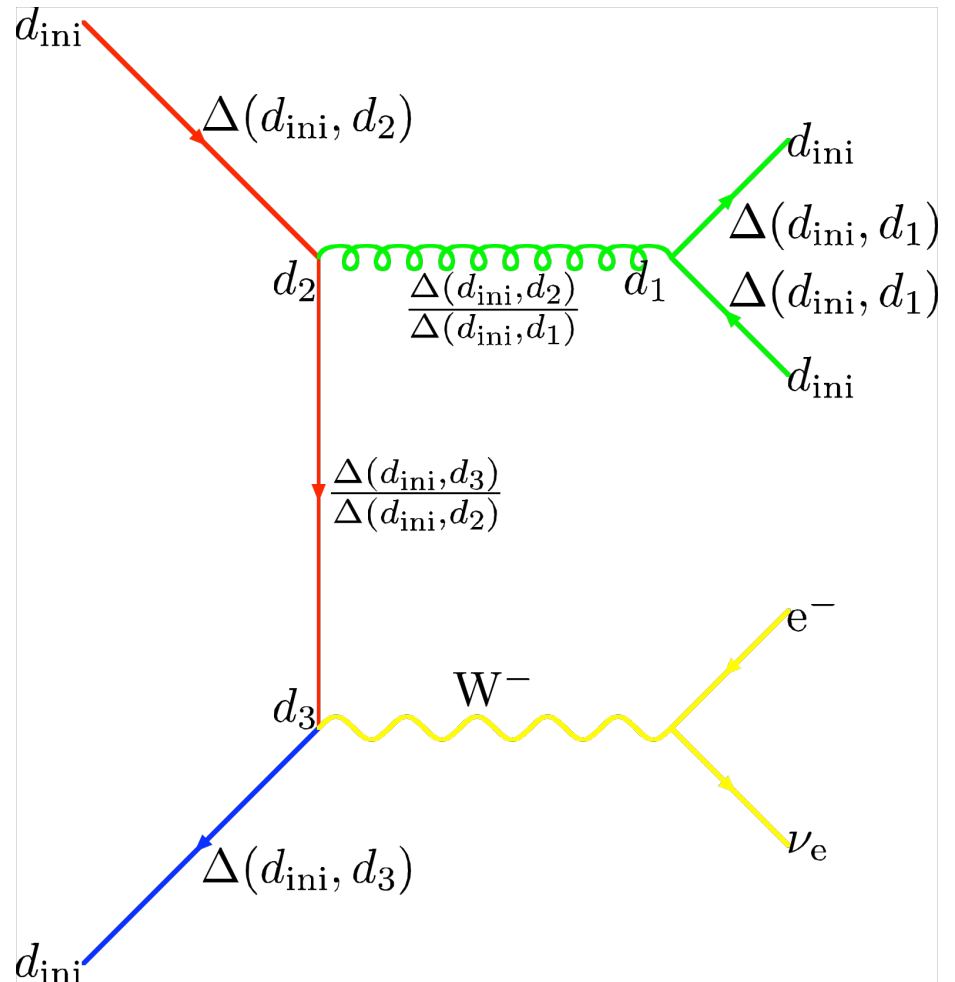


# CKKW Procedure

- 5) Reweight the lines by a Sudakov factor

$$\frac{\Delta(d_{\text{ini}}, d_j)}{\Delta(d_{\text{ini}}, d_k)}$$

- 6) Accept the configuration if the product of the  $\alpha_s$  and Sudakov weight is less than  $R \in [0,1]$  otherwise return to step 1.



# CKKW Procedure

- 7) Generate the parton shower from the event starting the evolution of each parton at the scale at which it was created and vetoing emission above the scale  $d_{\text{ini}}$ .



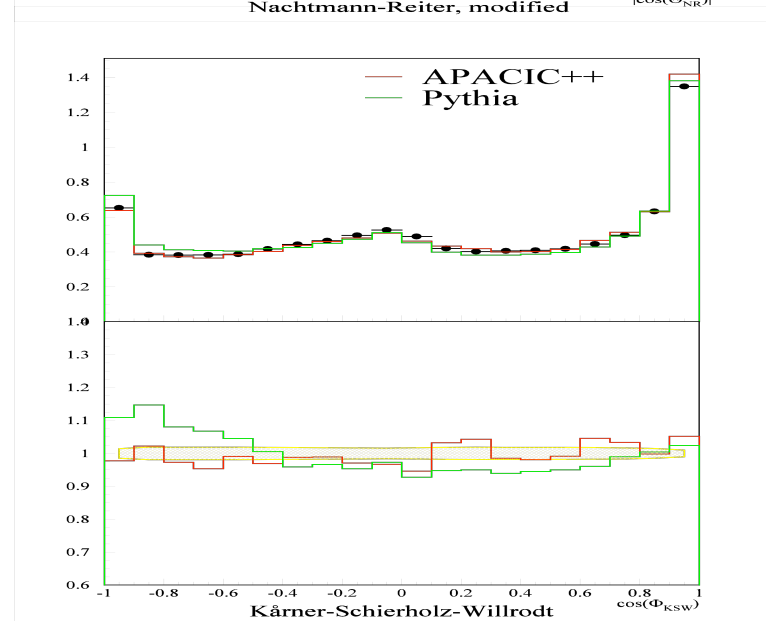
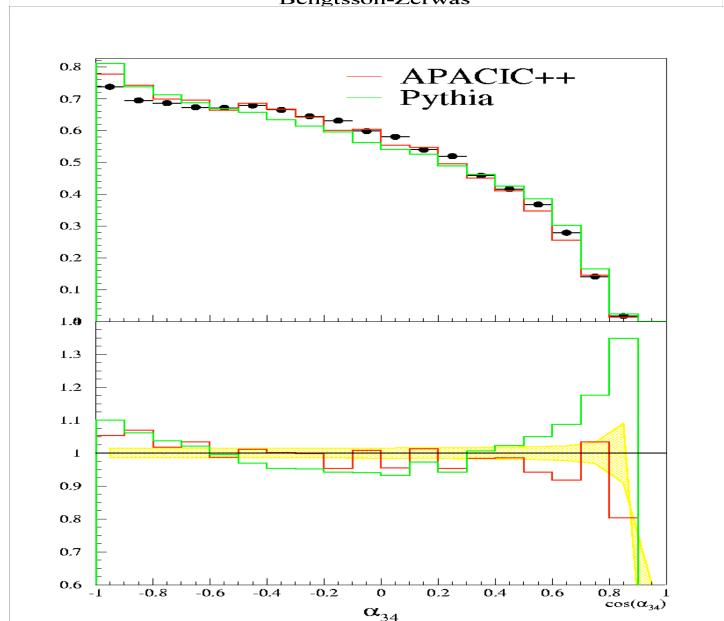
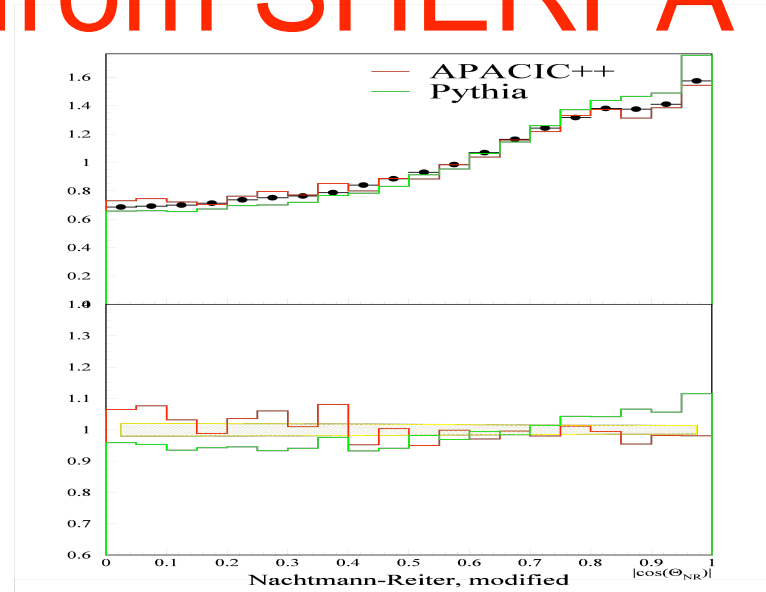
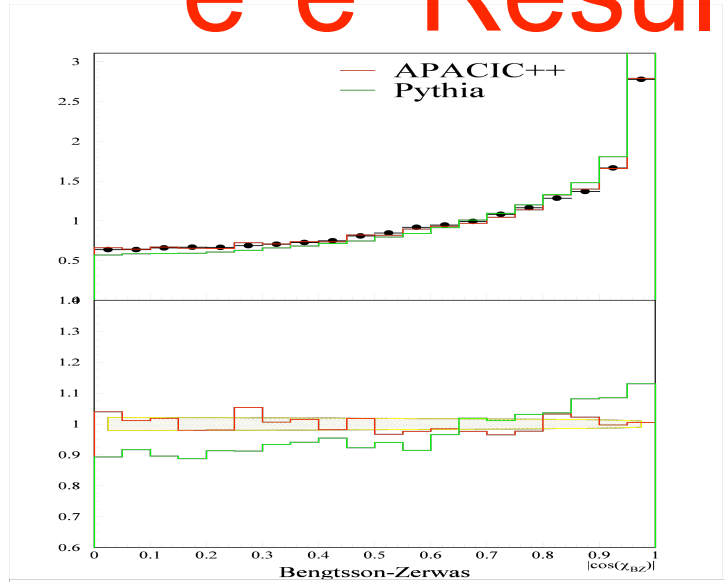
# CKKW Procedure

- Although this procedure ensures smooth matching at the NLL log level are **still choices** to be made:
  - Exact definition of the Sudakov form factors.
  - Scales in the strong coupling and  $\alpha_s$ .
  - Treatment of the highest Multiplicity matrix element.
  - Choice of the  $k_T$  algorithm.
- In practice the **problem** is **understanding** what the **shower** is doing and treating the matrix element in the same way.

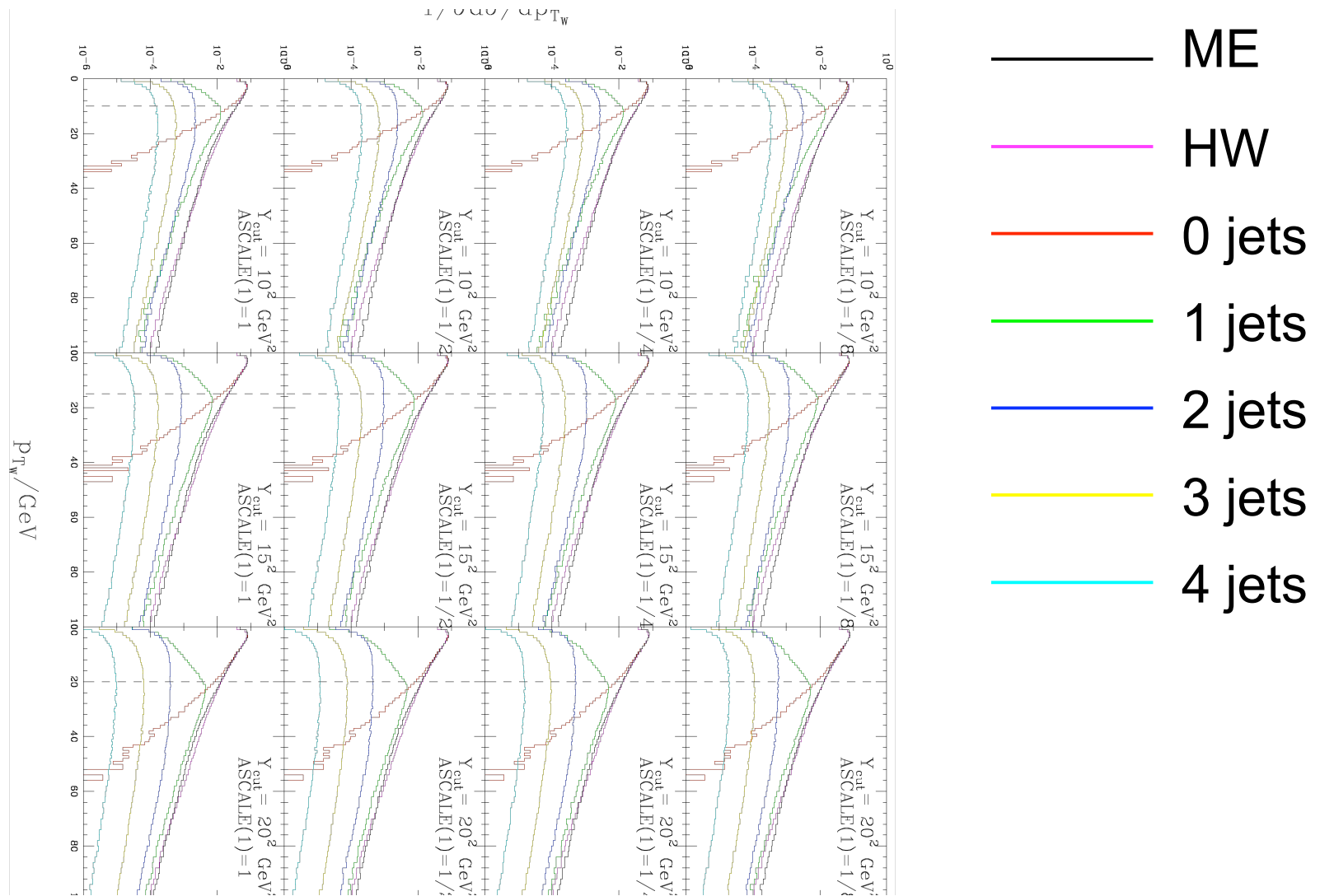
# CKKW Procedure

- A lot of work has been done mainly by
  - Frank Krauss et. al. (SHERPA)
  - Leif Lonnblad (ARIADNE)
  - Steve Mrenna (PYTHIA)
  - Peter Richardson (HERWIG)

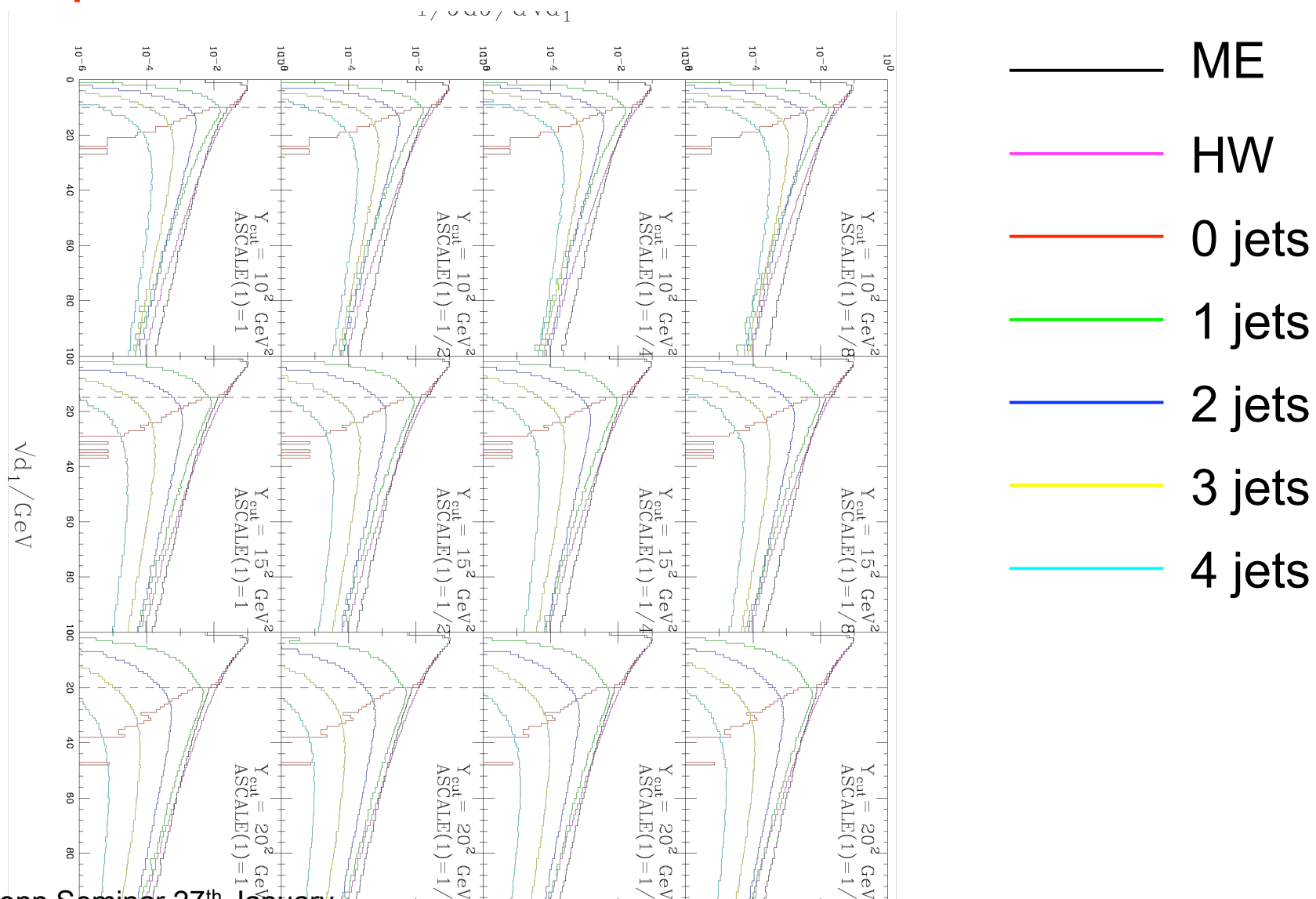
# $e^+e^-$ Results from SHERPA



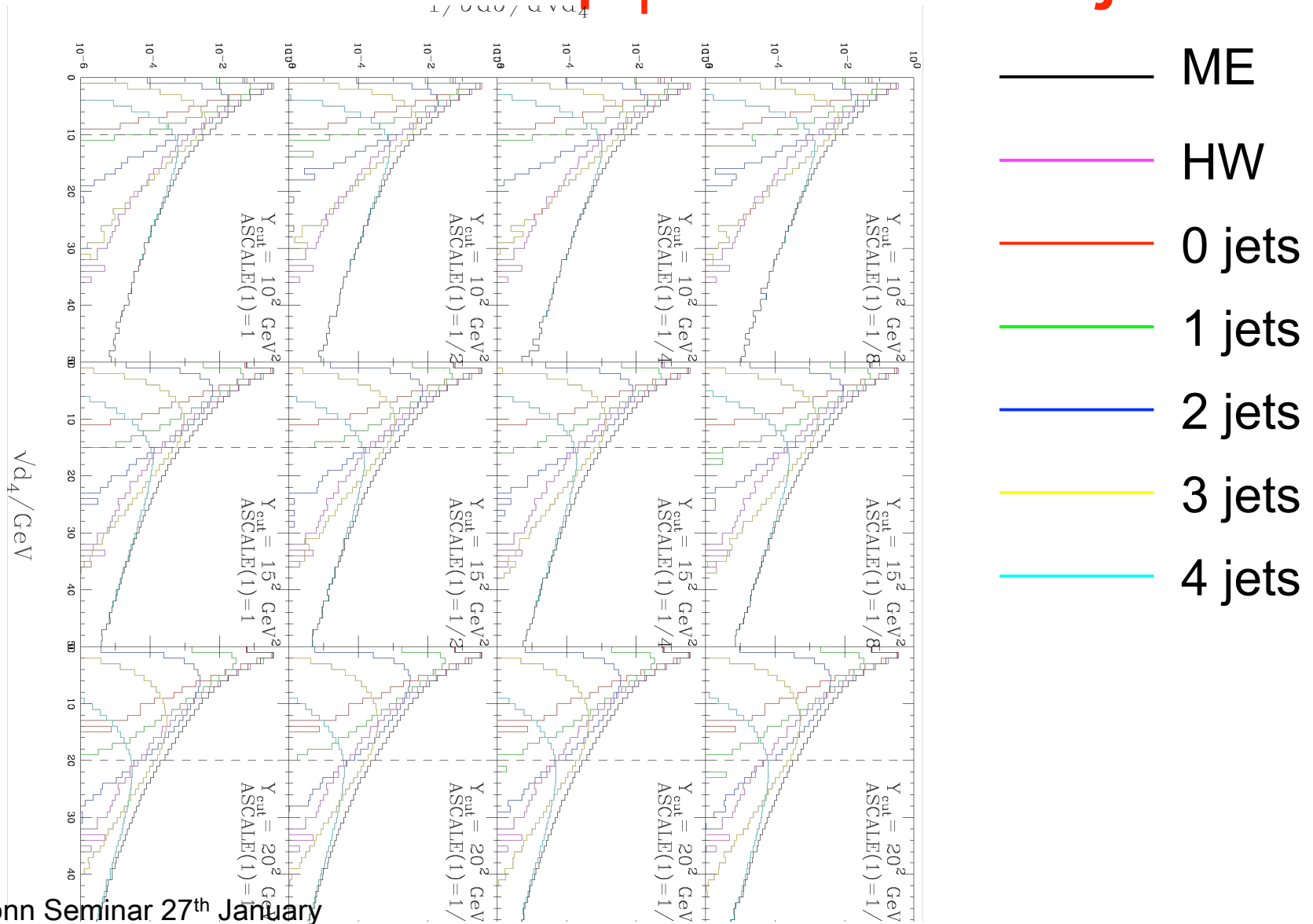
# $p_T$ of the W at the Tevatron



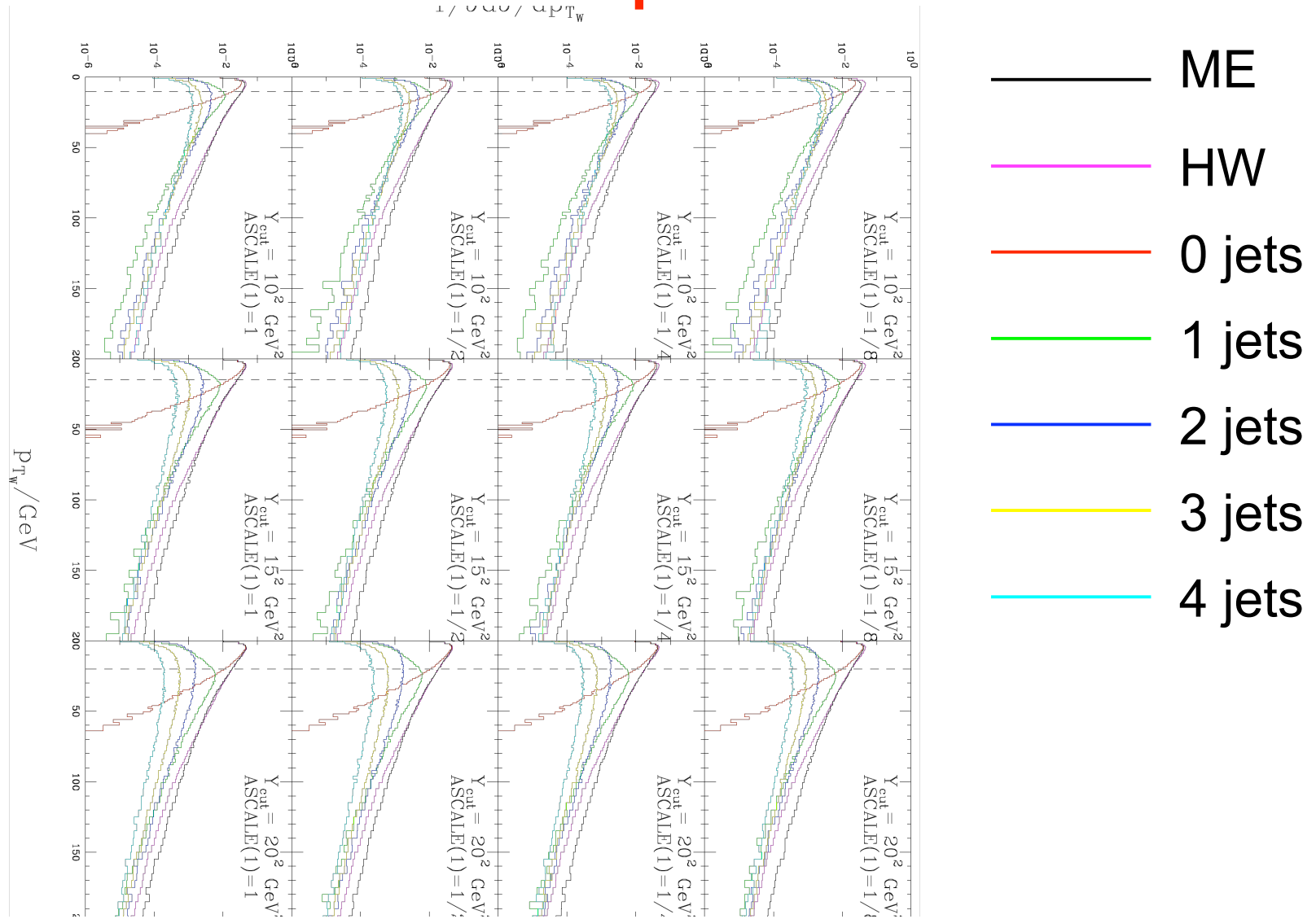
# $p_T$ of the hardest jet at the Tevatron



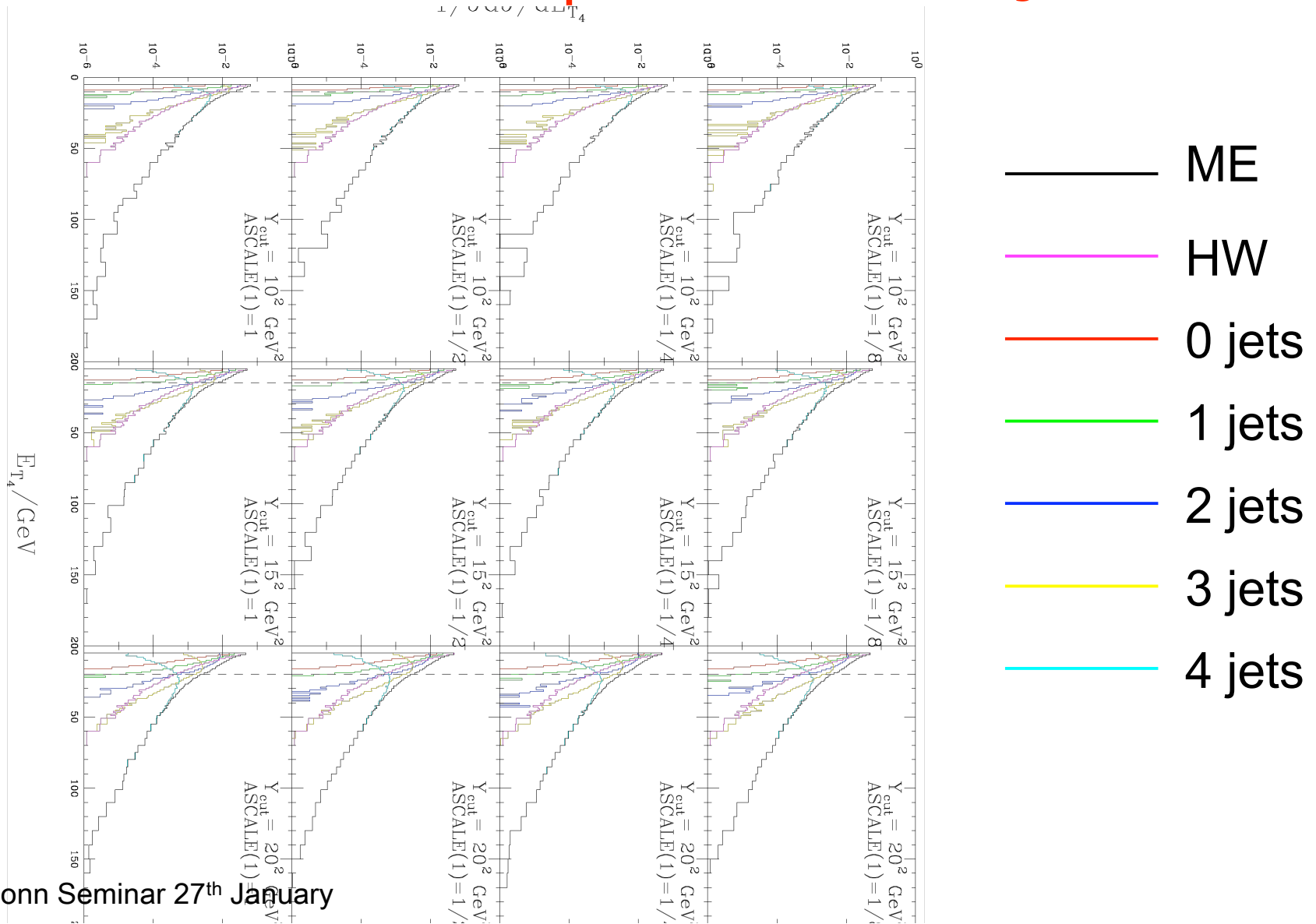
# Tevatron $p_T$ of the 4th jet



# LHC pt of W



# LHC $E_T$ of the 4th jet





# What Should I use?

- Hopefully this talk will help you decide which of the many different tools is most suitable for a given analysis.
  - Only soft jets relative to hard scale MC
  - Only one hard jet MC@NLO or old style ME correction
  - Many hard jets CKKW.
- The most important thing is to think first before running the simulation.

# Future

- Clearly much progress has been made with MC@NLO.
- The matching of many jets needs improved understanding of the shower and matching but is promising for many processes.
- Progress has been made with SHERPA.
- Hopefully the new Herwig++ and pT ordered PYTHIA shower's will have better properties for the matching.

# Future

- The Monte Carlo community is very small.
- There are three major projects
  - HERWIG (3 permanent staff, 3 postdocs, 1 student, ~3FTE)
  - PYTHIA (3 permanent staff, 1 postdoc, ~2FTE)
  - SHERPA (1 permanent staff, 4 students, ~4FTE)
- Given the large demand for both support and development this is not sustainable in the long term.
- We know how to construct the tools for the LHC.
- It may well be that everything we need will not be ready due to lack of manpower.